Journal
of
The Franklin Institute
Devoted to Science and the Mechanic Arts

Vol. 271 MAY, 1961 No. 5

CAN THERE BE A SHIELD FOR GRAVITATION? *

BY

W. F. G. SWANN [1]

I believe the high spot of interest of this organization has been the search for a "Shield for Gravitation." The origin of an expectation for a shield is obvious. We can shield our eyes from the sunlight by dark glasses. We can shield ourselves from the wind by a suitably placed screen. We can shield electrical apparatus from the effects of external electrical disturbances. By placing our portable radio inside a metal container, we can shield it so effectively from such external disturbances that it will not work at all. It will not even respond to the external influences which we wish it to respond to.

Our office is noisy. Every sound we make comes back to us by reflection from the walls, so we cover the walls with absorbing curtains or felt, and echoes are eliminated.

In an enlarged sense, we can shield our noses from unpleasant odors, from poisonous gases, or from the pollen which carries hay fever, by covering our nostrils with a suitable filtering appliance. We can shield our heads from the sun by wearing a hat.

And so, in the light of these experiences, we can well wonder whether it would be possible to reduce, or completely obliterate, the force of gravity by some kind of a shield. A falling body increases its velocity continually under the action of gravitation. Why should it not be possible to design, from some suitable material, a container whose walls

* This communication represents an amplification of an essay of the same title, submitted to the Gravity Research Foundation. The essay won one of the prizes offered by the Foundation and was presented verbally in the form here given at the Foundation's meeting held on August 27, 1960, at Gravity Village, New Boston, N. H.

[1] Director Emeritus, Bartol Research Foundation of The Franklin Institute, Swarthmore, Pa.; and a Senior Staff Advisor for The Franklin Institute Laboratories for Research and Development.

(Note—The Franklin Institute is not responsible for the statements and opinions advanced by contributors in the JOURNAL.)

would "absorb" the gravitational pull, so that the body inside would no longer fall?

Of course, during the last few years, we have become familiar with effects simulating rather closely those which would be caused by the reduction or complete elimination of the force of gravity. An animal traveling in one of the "sputniks" which encircle the earth experiences no sense of weight. If you put him in the center of his cabin without support, he would stay there, encircling the earth like a little satellite, quite independently of that other satellite which is his container. The phenomenon of weightlessness here presented to us is, however, not one involving a *shield* for gravitation, but one in which the animal in the sputnik is exposed to two oppositely directed influences, the force of gravity, and the effects resulting from the continual change in direction which he experiences as he travels around the earth. Indeed, a person who jumps off a stool experiences a sense of weightlessness as he falls, but it would not be correct to say that the force of gravity has been eliminated in the fall. We must rather hold that the force of gravity is present in its full intensity, but the acceleration he experiences during his fall produces upon all that responds to "force" within him an effect just equal and opposite to the effect produced by the force of gravitation, so that his condition is the same as if there were no force of gravitation and no acceleration.

If, indeed, we could find a means of shielding off gravitation, or even of reducing it, the effects on our everyday activities would be considerable. If we could reduce gravitation to the value it has on the moon, for example, a man could jump six times as high as he can at present. The limit of his horizontal jump, instead of being about 18 feet as it is at present, would be about 100 feet.

Now when, with the foregoing situation in mind, we approach a conventional man of science and ask him to produce a shield for gravitation, we find that he looks sad and shakes his head. If we ask him whether these unhappy and negative gestures imply that he "knows" that there can be no shield for gravitation, he will be compelled to admit that nobody can make such an assertion. If we ask him: "Why, then, do you look so miserable?" he will probably be driven to reply thus:

"My friend, in the development which science has experienced during the last few hundred years, I and my ancestors in science have learned certain things about the way in which the forces of nature act—the way in which electricity acts—the way in which gravitation acts, and the result of all this knowledge is such as to make me feel that a shield for gravitation would represent a very unnatural possibility. It is as though you asked me to produce, by cross fertilization, an animal which ran on wheels, asserting that such an achievement should be possible because automobiles run on wheels. I should have to tell you that the

whole structure of animals is such that I cannot see how a wheel can ever become a natural part of an animal's equipment. My knowledge of the necessity for blood circulation, and the difficulty of imagining any such process operative in the case of a wheel, puts that device into a category entirely different from those for which provision has been made in biology."

And if our man of science, inebriated with the exuberance of his own verbosity in giving us this answer, now looks a little less sad, and a little more companionable, it may be the time to ask him just what there is about the ordinary activities of nature which can look with equanimity upon the existence of a shield for certain forces, such as electrical forces, and yet compel us to look askance at a shield for gravitation. Having asked this question, let us make an attempt to answer it.

Now let me say at the outset that there *has* been a theory of gravitation which could, in principle, lend itself to the possibility of a shield for, or perhaps we had better say, an absorber of gravitational force. This theory of gravitation, put forward by Le Sage many years ago, postulated that all space was filled with small particles traveling in all directions with high speeds. Let us now consider two bodies, A and B, Fig. 1.

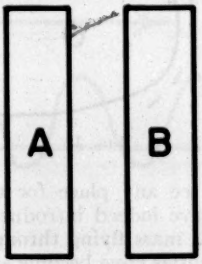

FIG. 1.

If A were alone, it would be bombarded by particles from all sides and would show no tendency to move. This would be true whether the particles falling upon it were totally or only partially absorbed. Now when B is present, some of the particles which would have come to A from the right are absorbed by B, so that A is bombarded more from the left than from the right and moves to the right. By analogous argument, A partially shields B against bombardment from the left, so that B tends to move to the left. The net result is that A and B move towards each other as though they attracted each other. If the bodies A and B were very thick in the left-right dimension, or if they were composed of some material very opaque to the particles, so that indeed

no particles reached the space between them, a body placed between them would show very little or no tendency to move towards the left or right as a result of the presence of external bodies to the left or right of both A and B, and we should have a "shield" for gravitation.

A hundred years ago, we might have accepted such a speculation as is here involved as something not unreasonable. For we recall the fact that light radiation from the sky is stopped completely by a very thin sheet of metal foil, whereas it can penetrate with ease the whole atmosphere which, as far as mass is concerned, represents the equivalent of 14 pounds per square inch, and we might well wonder whether there might be some substance so opaque to the flying particles that a relatively thin sheet of it would afford a shield. Alas for such speculations, however, we have learned so much about the structure of matter during the last century that we see no place for the existence of such a sub-

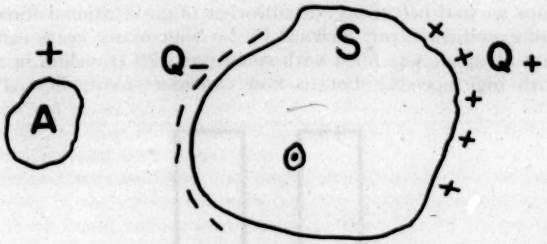

FIG. 2.

stance. Neither do we see any place for the flying particles. Our studies of cosmic rays have indeed introduced us to real particles of enormous energy per unit mass flying through space in all directions. However, our critical faculties have become so sharpened in the studies of these rays that, today, much more than formerly, we must look with considerable scepticism on the existence of such particles as would have the properties necessary to account for gravitation in the foregoing sense.

Let us, however, turn to the crucial differences between electric forces, on the one hand, and gravitational forces on the other hand, which make it so easy to understand shielding in the former realm and so difficult to understand it in the latter.

If A, Fig. 2, is a positively charged body, S a metallic shield, and O a point within it, we know that apparatus at O will be "shielded" from the influence of A. However, we do not visualize the shielding procedure as one in which the electrical forces emanating from A are somehow or other *absorbed* by S. We recognize the known fact of the existence of two opposite kinds of electricity, positive and negative. The

shield contains both kinds in equal amount, so that it is neutral as a whole. The positive electricity is fixed in the atoms of the "matter" of the shield, but in an electrical conductor the negative electrons are free to move, and they do so move as far as they can, toward A which attracts them, thus leaving a positive charge on the part of the shield remote from A. Now if we should put a charge of positive electricity at O, the force due to A would push it to the right. However, the push due to the charge labelled Q_+ and the pull due to the charge labelled Q_- combine to urge it to the left, It turns out that the push to the right resulting from A just balances the pushes and pulls due to Q_+ and Q_-, respectively, and so the charge at O is "shielded" from the external influence originating at A. We see that in this case the ultimate phenomenon of shielding depends upon the existence of two opposite kinds of electricity. If there were two kinds of gravitational matter, one kind K_1 which repelled its kind, and another kind K_2, which repelled *its* kind, and if the kind K_1 attracted the kind K_2, then we might hope to realize shields for gravitation in the same manner as we realize shields for electrical forces.

Consider another rather more subtle case, where a beam of light waves or radio waves impinges upon a plate P, Fig. 3, with the result

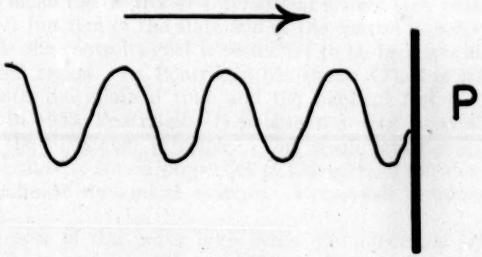

FIG. 3.

that nothing is observed to the right of P. The accepted way of regarding this phenomenon is not to suppose that the waves are somehow or other absorbed by P. We, indeed, say that *nothing* happens to the waves. They go on right through P, off to infinity. If we ask: "Why then do we not observe them?," the answer is that they set up electric currents in the plate P, which currents send out waves on their own to the right of P, but waves which are "out of step" with the original set in the sense that the "crests" of one set coincide with the "troughs" of the other. Thus, while there is "nothing" to the right of P, the nothing is, as it were, composed of equal and opposite halves. If P is what we call a complete reflector, the electric currents set up in it send waves

to the left as well as to the right and the nature of things is such that the waves which go to the right are out of step with the original set, and so cancel them, while the waves which go to the left cooperate with the original set to produce what we call reflection. If gravitational forces could be explained in terms of waves, and if matter were such that these waves, when impinging upon suitable matter, set up other gravitational waves, we might be able to realize a shield for gravitation in the same way that we shielded the radio, introduced at the beginning of this essay, against the waves from without by the metal container with which we surrounded it.

PONTRIAGIN'S MAXIMUM PRINCIPLE AND THE PRINCIPLE OF OPTIMALITY *

BY

CHARLES A. DESOER [1]

ABSTRACT

This paper demonstrates that Pontriagin's Maximum Principle may be derived from the principle of Optimality. It considers a control system described by $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, t)$ where the control vector $\mathbf{u}$ is restricted to a closed and bounded set. The optimal control steers the system from an initial state $\mathbf{x}_0$ at t_0 to a moving target in such a way that the cost of control along the optimal trajectory $\int_{t_0}^{t_1} f_0[\mathbf{x}^0(\tau), \mathbf{u}^0(\tau), \tau]d\tau$ is minimized. The optimal control $\mathbf{u}^0$ satisfies the maximum principle at each point of the trajectory: $\max_{\mathbf{u},\Omega} \langle \tilde{g}(\mathbf{x}), \tilde{f}(\mathbf{x}, \mathbf{u}) \rangle = \langle \tilde{g}(x), \tilde{f}(\mathbf{x}, \mathbf{u}^0) \rangle$ where the vector $\tilde{g}(\mathbf{x})$ is related to $\nabla\Gamma$ where $\Gamma(\mathbf{x}_0, \mathbf{x})$ is the minimal cost of going from $\mathbf{x}_0$ to $\mathbf{x}$.

I. INTRODUCTION

Interest in optimal control problems has been growing at a very fast rate both in this country and in the Soviet Union (1–10).[2] By optimal control we mean the control of a dynamical system such that the time integral of a function of the state and of the control signal(s) is minimized while the control signal is restricted to be in a specified closed and bounded region. As Pontriagin has shown (7), this formulation includes both the minimal time and the minimal fuel consumption problem. In 1959, Pontriagin (7) published a very interesting result known as the maximum principle. This result was particularly interesting because it stated properties of the optimal solution for very general nonlinear dynamical systems. Krasovskii obtained similar results (10).

The purpose of this paper is to derive the maximum principle of Pontriagin in a completely different fashion. The point of view adopted here is to use the principle of optimality. As shown recently by S. E. Dreyfus (11), this point of view is particularly powerful to obtain the necessary conditions satisfied by the solutions of variational problems. As will become apparent in the following, the derivation is particularly simple and, in particular, does not require Ω to be convex. In fact our derivation will make it obvious that the validity of the maximum principle does not depend on the terminal conditions imposed on the optimal trajectory, that is, the maximum principle is valid if the opti-

* The research reported in this paper has been sponsored by the Electronics Research Directorate of the Air Force Cambridge Research Center, Air Research and Development Command, under contract AF 19(604)-5466.

[1] Department of Electrical Engineering, University of California, Berkeley, Calif.

[2] The boldface numbers in parentheses refer to the references appended to this paper.

mal trajectory must reach a fixed point in state space, a moving point in state space (with a specified law of motion), or a moving subset of the state space. In their latest paper (8), Pontriagin *et al.* reached the same conclusion by extending their previous method.

II. STATEMENT OF THE PROBLEM

We are considering a problem more general than that considered by Pontriagin (7). More precisely we consider a nonautonomous nonlinear dynamic system S which can be controlled by r control signals $u_1, u_2, \cdots, u_r$. The system is described by the law of motion of its state vector $\mathbf{x}$ in n-dimensional state space R^n:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, t) \tag{1}$$

where

(*a*) $\mathbf{x} = \text{col}(x_1, x_2, \cdots x_n)$ is the state vector in the n-dimensional space R^n;

(*b*) the control vector $\mathbf{u}$ is restricted to remain in the closed and bounded set Ω; for example, Ω might be the parallelepiped $\{\mathbf{u}: |u_i| \leqslant \gamma_i, i = 1, 2, \cdots r\}$;

(*c*) the vector valued function $\mathbf{f}(\mathbf{x}, \mathbf{u}, t)$ is assumed to have continuous first partial derivatives at all points of $R^n \times \Omega \times I$, I being the real time axis.

Any control function $\mathbf{u}(t)$ that satisfies condition (*b*) is said to be *admissible*. Suppose that $\mathbf{x}$, the state vector of the system S, represents the motion of an object in physical three-dimensional space. Let then x_1, x_2, x_3 be the Cartesian coordinates of its center of gravity in that space. The problem we are considering involves a target T represented by its law of motion in the physical three-dimensional space: $z_1(t), z_2(t), z_3(t)$, three specified functions of time. The problem is to steer the system S from the specified starting point $\mathbf{x}_0, t_0$ until it hits the target T (in physical three-space) while minimizing the integral

$$\int_{t_0}^{t_1} f_0[\mathbf{x}^0(\tau), \mathbf{u}^0(\tau), \tau]d\tau \tag{2}$$

where

(1) f_0 is a specified cost function,

(2) $\mathbf{x}^0(t)$ and $\mathbf{u}^0(t)$ are the optimal trajectory and the optimal control,

(3) t_1 is the instant of collision, that is, the instant at which $x_i(t_1) = z_i(t_1)$ for $i = 1, 2, 3$. Observe that t_1 is not specified in advance.

Before reformulating completely the problem in terms of concepts involving only the space R^n, observe that we may, without loss of generality, assume that f and f_0, of (1) and (2), respectively, do not depend explicitly on time: indeed if they were dependent explicitly on

time we would introduce an $(n+1)^{\text{th}}$ state variable $x_{n+1} = t$ and an $(n+1)^{\text{th}}$ differential equation $\dot{x}_{n+1} = 1$. Therefore, from now on we replace (1) and (2) by (1a) and (2a) respectively:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}) \qquad \mathbf{x} \epsilon R^n, \ \mathbf{u} \epsilon \Omega \tag{1a}$$

$$\int_{t_0}^{t_1} f_0(\mathbf{x}^0(\tau), \mathbf{u}^0(\tau)) d\tau. \tag{2a}$$

The target motion in physical three space suggests the consideration of the cylinder C in R^n: $C = \{\mathbf{x}(t) \mid x_i(t) = z_i(t) \text{ for } i = 1, 2, 3; x_j \text{ arbitrary for } j = 4, \cdots, n\}$. In general, C is a subset of R^n that moves with time.

The problem can be restated in R^n as follows: find an admissible control $\mathbf{u}(t)$, $t \epsilon [t_0, t_1]$ which will bring the system S from $\mathbf{x}_0$, t_0 on the cylinder C so that the integral (2a) is minimized.

III. THE MAXIMUM PRINCIPLE

Assume that an optimal trajectory exists. Let it be represented by $\mathbf{x}^0(t)$, $t \epsilon [t_0, t_1]$. Consider an arbitrary point $\mathbf{x}^0(t)$ on the optimal trajectory, with $t_0 < t < t_1$. For convenience we write henceforth $\mathbf{x}$ for $\mathbf{x}^0(t)$. The principle of optimality states that an optimal control has the property that whatever the initial state and the initial decision are, the remaining decisions must constitute an optimal control with regard to the state resulting from the first decision. Let us express this fact for a special class of other possible trajectories. Let Δ be a small positive interval of time. Consider all admissible controls $\mathbf{u}$ over the interval $(t - \Delta, t)$. From (1a), any trajectory using such a control will go through the point $\mathbf{x} - \mathbf{f}(\mathbf{x}, \mathbf{u})\Delta + o(\Delta)$. If $\Gamma(\mathbf{x}_0, \mathbf{x})$ is the minimal cost of going from $\mathbf{x}_0$ to $\mathbf{x}$ (that is, the cost computed along the optimal trajectory), then the cost along the perturbed trajectory just mentioned is

$$\Gamma(\mathbf{x}_0, \mathbf{x} - \mathbf{f}(\mathbf{x}, \mathbf{u})\Delta + o(\Delta)) + f_0(\mathbf{x}, \mathbf{u})\Delta + o(\Delta)$$

where the last two terms give the contribution to the cost given by (2a) during the interval $(t - \Delta, t)$. Therefore the minimal cost of going from $\mathbf{x}_0$ to $\mathbf{x}$ is characterized by

$$\Gamma(\mathbf{x}_0, \mathbf{x}) = \min_{\mathbf{u}(t) \epsilon \Omega} \{\Gamma(\mathbf{x}_0, \mathbf{x} - \mathbf{f}(\mathbf{x}, \mathbf{u})\Delta + o(\Delta)] + f_0(\mathbf{x}, \mathbf{u})\Delta\} + o(\Delta)$$

which expresses the optimality of the control $\mathbf{u}$ over the interval $(t - \Delta, t)$. Now if $\Gamma(\mathbf{x}_0, \mathbf{x})$ has continuous first derivatives in an open set containing $\mathbf{x}$, we may use the mean value theorem for functions of several variables (12) and in the limit of $\Delta \to 0$ we get, after cancelling

terms and changing signs,

$$0 = \max_{u \in \Omega} \{ \nabla \Gamma(\mathbf{x}_0, \mathbf{x}) \cdot \mathbf{f}(\mathbf{x}, \mathbf{u}) - f_0(\mathbf{x}, \mathbf{u}) \} \tag{3}$$

where $\nabla \Gamma(\mathbf{x}_0, \mathbf{x})$ is the gradient of the scalar $\Gamma(\mathbf{x}_0, \mathbf{x})$ evaluated at $\mathbf{x}$. In order to write condition (3) more elegantly, introduce the following $n + 1$ dimensional vectors

$$\tilde{f}(\mathbf{x}, \mathbf{u}) = \text{col } (f_0(\mathbf{x}, \mathbf{u}), f_1(\mathbf{x}, \mathbf{u}), \cdots, f_n(\mathbf{x}, \mathbf{u})) \tag{4}$$

$$\tilde{g}(x) = \text{col } \left(-1, \frac{\partial \Gamma}{\partial x_1}, \frac{\partial \Gamma}{\partial x_2}, \cdots, \frac{\partial \Gamma}{\partial x_n} \right). \tag{5}$$

Hence, using the notation $\langle \mathbf{a}, \mathbf{b} \rangle$ to indicate the scalar product of $\mathbf{a}$ and $\mathbf{b}$,

$$\max_{u \in \Omega} \langle \tilde{g}(\mathbf{x}), \tilde{f}(x, u) \rangle = 0. \tag{6}$$

This is the maximum principle. Its meaning is the following: if there is an optimal trajectory $\mathbf{x}^0(t)$ (and an associated optimal control $\mathbf{u}^0(t)$) and if $\Gamma(\mathbf{x}_0 \mathbf{x}) \epsilon C^1$ at $\mathbf{x}$, then at each point $\mathbf{x}$ of the trajectory the optimal control $\mathbf{u}^0$ maximizes the scalar product in (6), that is,

$$\max_{u \in \Omega} \langle \tilde{g}(\mathbf{x}), \tilde{f}(\mathbf{x}, \mathbf{u}) \rangle = \langle \tilde{g}(\mathbf{x}), \tilde{f}(\mathbf{x}, \mathbf{u}^0) \rangle. \tag{7}$$

Furthermore, the maximum value of this scalar product is 0,

$$\langle \tilde{g}(\mathbf{x}), \tilde{f}(x, u^0) \rangle = 0. \tag{8}$$

IV. THE DIFFERENTIAL EQUATIONS SATISFIED BY $\tilde{g}(x)$

Consider the successive values taken by $\tilde{g}(\mathbf{x})$ along the optimal trajectory. We shall write $\tilde{g}(t)$ for $\tilde{g}[\mathbf{x}^0(t)]$. For $i = 1, 2, \cdots, n$, we have, assuming that Γ has continuous second partial derivatives (12),

$$\frac{dg_1}{dt} = \frac{d}{dt} \frac{\partial \Gamma}{\partial x_i} = \sum_{j=1}^{n} \frac{\partial^2 \Gamma}{\partial x_j \partial x_i} x_j^0 = \sum_{j=1}^{n} \frac{\partial^2 \Gamma}{\partial x_i \partial x_j} f_j(\mathbf{x}, \mathbf{u}^0)$$

and

$$\frac{dg_0}{dt} = 0,$$

since $g_0 = -1$ everywhere. Now, observe that (8) defines $\mathbf{u}^0$ implicitly for each $\mathbf{x}$, that is, it provides a means for computing $\mathbf{u}^0$ at each point $\mathbf{x}$. Therefore as we consider (8) for various values of $\mathbf{x}$, the corresponding value of $\mathbf{u}^0$ varies with $\mathbf{x}$ and if we differentiate (8) with respect to

x_i we have

$$\sum_{j=0}^{n} \frac{\partial g_j}{\partial x_i} f_j(x, \mathbf{u}^0) = -\sum_{j=0}^{n} g_j \left(\frac{\partial f_j}{\partial x_i} + \sum_{l=1}^{r} \frac{\partial f_j}{\partial u_l} \frac{\partial u_l^0}{\partial x_i} \right).$$

Hence the differential equations for the g_i's become

$$\frac{dg_i}{dt} = -\sum_{j=0}^{n} g_j \left(\frac{\partial f_j(\mathbf{x}, \mathbf{u}^0)}{\partial x_i} + \sum_{l=1}^{r} \frac{\partial f_j(\mathbf{x}, \mathbf{u}^0)}{\partial u_l} \frac{\partial u_l^0(\mathbf{x})}{\partial x_i} \right) \quad (i = 1, 2, \cdots, n)$$

$$\frac{dg_0}{dt} = 0 \qquad g_0(t_0) = -1. \tag{9}$$

Observe that if the maximizing $\mathbf{u}$, namely, $\mathbf{u}^0$ of (7) actually is in the interior of Ω, then

$$\sum_{j=0}^{n} g_j \frac{\partial f_j}{\partial u_l} = 0 \qquad l = 1, 2, \cdots, r$$

and (9) simplifies to

$$\frac{dg_i}{dt} = -\sum_{j=0}^{n} g_j \frac{\partial f_j(\mathbf{x}, \mathbf{u}^0)}{\partial x_i}.$$

If, however, $\mathbf{u}^0$ lies on the boundary of Ω, this simplification will not occur. It is interesting to note that Eqs. 9 form the adjoint system to the first variation of (1a), provided in (1a) $\mathbf{u}^0$ is considered as a function of $\mathbf{x}$.

V. REMARKS

Equation 6 is a statement of Pontriagin's maximum principle (7). Our derivation, however, is completely different and shows that the maximum principle holds irrespective of the terminal conditions imposed on the optimal trajectory.

The vector $\bar{\mathbf{g}}$ satisfies also the adjoint system to the first variation of (1a), as does the vector ψ in Ref. 7. However, in the present derivation $\mathbf{u}^0$ appears as a function of $\mathbf{x}$ and, therefore, must contribute to the first variation.

The above derivation could be repeated, with obvious modifications, if the minimal cost $C(\mathbf{x}, \mathbf{x}_1)$, of going from $\mathbf{x}$ to $\mathbf{x}_1$ were used instead of $\Gamma(\mathbf{x}_0, \mathbf{x})$. In general, $\nabla\Gamma$ and ∇C bear no simple relation to each other because it is only along the optimal trajectory that $\Gamma(\mathbf{x}_0, \mathbf{x}) + C(\mathbf{x}, \mathbf{x}_1) = \Gamma(\mathbf{x}_0, \mathbf{x}_1)$. Examples show that in cases where (3) does not specify $\mathbf{u}^0$ uniquely, the corresponding relation (with $\nabla\Gamma$ replaced by $-\nabla C$) may determine $\mathbf{u}^0$ uniquely.

In the special case of a minimal time problem, that is, $f_0(\mathbf{x}, \mathbf{u}) = 1$ on $R^n \times \Omega$, the maximum principle reduces to

$$\max_{u \in \Omega} \langle \nabla\Gamma(\mathbf{x}_0, \mathbf{x}), \mathbf{f}(\mathbf{x}, \mathbf{u}) \rangle = \langle \nabla\Gamma(\mathbf{x}_0, \mathbf{x}), \mathbf{f}(\mathbf{x}, \mathbf{u}^0) \rangle = 1.$$

Physically, this is obvious: the middle member of the above equality is the rate at which time increases along the optimal trajectory; hence should always be equal to 1; the first member states that the optimal control $\mathbf{u}^0$ should be such that at each point $\mathbf{x}$, the optimal instantaneous velocity $f(\mathbf{x}, \mathbf{u}^0)$ should be directed so that the system reaches points furthest in time (as measured by $\Gamma(\mathbf{x}_0, \mathbf{x})$) as possible.

In some cases the maximum principle points to significant properties of the optimal control. Consider, for example, a nonlinear autonomous system

$$\dot{\mathbf{x}} = \boldsymbol{\phi}(\mathbf{x}) + \mathbf{B}\mathbf{u} \qquad \mathbf{u} \in \Omega$$

where

(*a*) $\boldsymbol{\phi}(\mathbf{x})$ is a vector function of $\mathbf{x}$, usually nonlinear, satisfying a Lipschitz condition for all $\mathbf{x}$.

(*b*) $\mathbf{B}$ is an $n \times r$ matrix.

(*c*) Ω is the parallelepiped $|u_i| \leqslant \gamma_i \qquad i = 1, 2, \cdots, r$

(*d*) $f_0(\mathbf{x}, \mathbf{u}) = 1$ on $R^n \times r$, that is, we consider a minimal time problem. Then the maximum principle states that at each point of the optimal trajectory, the optimal control maximizes

$$\langle \mathbf{g}, \mathbf{B}\mathbf{u} \rangle = \langle \mathbf{B}'\mathbf{g}, \mathbf{u} \rangle$$

hence

$$u_i = \operatorname{sgn} \left(\sum_{j=1}^{n} b_{ji} g_j \right) \qquad i = 1, 2, \cdots, r$$

in other words, the optimal control for these nonlinear systems will be of the bang-bang type.

VI. CONCLUSION

We have shown that Pontriagin's maximum principle simply expresses the optimality of the choice of control function at each point of the trajectory. The present derivation makes no assumption on the set Ω except that it is closed and bounded. In fact, the whole derivation shows that the maximum principle will always be satisfied irrespective of the particular terminal conditions imposed on the optimal trajectory.

It should be kept in mind that the conditions derived here are *necessary*, that is, if an optimal solution exists and if $\mathbf{f}$ and Γ have continuous first derivatives, then the maximum principle holds at each point of the optimal trajectory.

REFERENCES

(1) D. C. McDONALD, "Nonlinear Techniques for Improving Servo Performance," *Proc. Natl. Electronics Conf.*, Vol. 6, pp. 400–421 (1950).

(2) R. W. BASS, "Equivalent Linearization, Nonlinear Circuit Synthesis and the Stabilization and Optimization of Control Systems," *Proc. 2nd Nonlinear Circuit Analysis Symposium*, pp. 163–198, New York Polytechnic Institute of Brooklyn, 1956.

(3) D. W. BUSHAW, "Optimal Discontinuous Forcing Terms," in "Contributions to the Theory of Nonlinear Oscillations," edited by S. Lefshetz, Vol. IV, Princeton, Princeton Univ. Press, 1958, pp. 29–52.

(4) R. BELLMAN, I. GLICKSBERG AND O. GROSS, "On the Bang Bang Control Problem," *Quart. J. Appl. Math.*, Vol. 14, pp. 11–18 (1956).

(5) C. A. DESOER, "The Bang Bang Servo Problem Treated by Variational Techniques," *Information and Control*, Vol. 2, pp. 33–348 (1959).

(6) E. B. LEE, "Mathematical Aspects of the Synthesis of Linear Minimum Response-Time Controllers," *IRE Trans. PGAC AC-5*, Vol. 4, pp. 283–284 (1960).

(7) L. S. PONTRIAGIN, "Optimal Control Processes," *Uspekhi Matem. Nauk*, Vol. 14, No. 1, pp. 3–20 (1959). (In English: *Automation Express*, Vol. 1, Nos. 10 and 11.)

(8) V. G. BOLTYANSKII, R. V. GAMKRELIDZE AND L. S. PONTRIAGIN, "Optimal Control Processes," *Izvestia Akad. Nauk SSSR*, Ser. Mat., Vol. 24, No. 1, pp. 3–40 (1960).

(9) L. I. ROZONOER, "The Maximum Principle of L. S. Pontriagin in Optimal System Theory," *Automation and Remote Control*, Vol. 20, Nos. 10, 11, 12 (Oct.–Dec. 1959).

(10) N. N. KRASOVSKII, "On a Problem of Optimum Control of Nonlinear Systems," *P.M.M.*, Vol. 23, No. 2, pp. 209–229 (1959).

(11) S. E. DREYFUS, "Dynamic Programming and the Calculus of Variation," *J. Math. Anal. and Appl.*, Vol. I, pp. 228–239 (1960).

(12) T. M. APOSTOL, "Mathematical Analysis," Reading (Mass.), Addison-Wesley Publishing Co., Theorems 6-17 and 6-18, 1957.

A NOTE ON THE RELATIVISTIC BOLTZMANN EQUATION AND SOME APPLICATIONS

BY

S. YADAVALLI [1]

ABSTRACT

Starting from the Boltzmann equation in the proper frame, the relativistic version of the Boltzmann equation in the absence of collisions is derived. As an application, a simple example of a velocity modulated drifting relativistic electron beam in the presence of a velocity distribution is treated.

I. INTRODUCTION

In the solution of many problems in plasma physics, the kinetic method involving the Boltzmann equation is often employed. The standard nonrelativistic form of the Boltzmann equation is well known (1).[2] Relativistic versions of the Boltzmann equation have been either given or written in the past by others (2, 3, 4, 5). An alternative derivation of the Boltzmann equation in the absence of collision is given below, and the resulting equation is expressed in a form identical with that of Clemmow and Willson (2) for purposes of comparison. A simple example of a velocity modulated drifting electron beam is treated below.

II. DERIVATION OF THE RELATIVISTIC BOLTZMANN EQUATION

Let $f(\vec{x}, \vec{v}, t)$ be the one particle distribution function in phase space $\vec{x}$, $\vec{v}$ where $\vec{x}$ and $\vec{v}$ are the ordinary spatial and velocity coordinates and t the time. Then, the Boltzmann equation in the absence of collisions may be written as

$$\frac{df}{dt} = 0. \tag{1}$$

Equation 1 is the standard non-relativistic form of the Boltzmann equation in the absence of collisions (1).

In the relativistic case, one may start with the Boltzmann equation in the proper frame by writing

$$\frac{df}{ds} = 0 \tag{2}$$

[1] Stanford Research Institute, Menlo Park, Calif.

[2] The boldface numbers in parentheses refer to the references appended to this paper.

and transform the equation to any other frame of interest, as done below. In Eq. 2, ds is the proper time interval times c (the speed of light), that is

$$ds = c\,dt\left(1 - \frac{v^2}{c^2}\right)^{1/2}.$$

Equation 2 can also be written as

$$\frac{df}{ds} = \frac{\partial f}{\partial x_j}\frac{dx_j}{ds} + \frac{\partial f}{\partial u_j}\frac{du_j}{ds} = 0 \tag{3}$$

where the summation convention of repeated indices is understood. In the four vector notation (6)

$$x_j = (x, y, z, ict) = (x, y, z, \tau)$$

$$u_j = \frac{dx_j}{ds}$$

$$u_\alpha = \frac{v_\alpha}{c\left(1 - \frac{v^2}{c^2}\right)^{1/2}}$$

$$u_4 = \frac{i}{\left(1 - \frac{v^2}{c^2}\right)^{1/2}}$$

where Latin indices designate the four vectors, and Greek indices denote ordinary vectors; also $\vec{v}$ is the ordinary velocity and $i = \sqrt{-1}$. Note that: $j = 1, 2, 3$, and 4, and $\alpha = 1, 2, 3$. One also has

$$w_j = \frac{du_j}{ds}$$

$$w_\alpha = \frac{1}{c^2\left(1 - \frac{v^2}{c^2}\right)^{1/2}}\frac{d}{dt}\left(\frac{v_\alpha}{\left(1 - \frac{v^2}{c^2}\right)^{1/2}}\right)$$

$$w_4 = \frac{i}{c\left(1 - \frac{v^2}{c^2}\right)^{1/2}}\frac{d}{dt}\left(\frac{1}{\left(1 - \frac{v^2}{c^2}\right)^{1/2}}\right)$$

$$(dx_j)^2 = -\,ds^2$$

$$u_j^2 = -\,1$$

$$u_j w_j = 0.$$

From the above relations, one notes that

$$\frac{\partial f}{\partial x_j} = \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3} \quad \text{for} \quad j = 1, 2, 3$$

$$\frac{\partial f}{\partial x_j} = \frac{1}{c}\frac{\partial f}{\partial t} \quad \text{for} \quad j = 4$$

$$\frac{dx_j}{ds} = \frac{v_1}{c\left(1 - \frac{v^2}{c^2}\right)^{1/2}}, \frac{v_2}{c\left(1 - \frac{v^2}{c^2}\right)^{1/2}}, \frac{v_3}{c\left(1 - \frac{v^2}{c^2}\right)^{1/2}} \quad \text{for} \quad j = 1, 2, 3 \tag{4}$$

$$\frac{dx_j}{ds} = \frac{i}{\left(1 - \frac{v^2}{c^2}\right)^{1/2}} \quad \text{for} \quad j = 4.$$

Substitution of relations (4) in the first term of Eq. 3 yields

$$\frac{\partial f}{\partial x_j}\frac{dx_j}{ds} = \frac{1}{c\left(1 - \frac{v^2}{c^2}\right)^{1/2}}\left\{\frac{\partial f}{\partial t} + \vec{v}\cdot\vec{\nabla}_r f\right\} \tag{5}$$

$\vec{\nabla}_r$ denotes the spatial gradient.

At this point we note the following. The distribution function f is a function of v_1, v_2, v_3; f can also be expressed as a function of any three components of the four vector u_j because conservation of energy and momentum are assured. Since it is desired to show the equivalence of this with the equation obtained in (2), we assume that the one particle distribution function is a function of v_1, v_2, and v_3.

From the above transformation relations, one has

$$\left.\begin{aligned} \frac{\partial f}{\partial u_\alpha} &= \frac{\partial f}{\partial v_\alpha} c\left(1 - \frac{v^2}{c^2}\right)^{3/2} \\ \frac{\partial f}{\partial u_\alpha}\frac{du_\alpha}{ds} &= \frac{\partial f}{\partial v_\alpha} c\left(1 - \frac{v^2}{c^2}\right)^{3/2} \cdot \frac{1}{c^2\left(1 - \frac{v^2}{c^2}\right)^{1/2}} \frac{d}{dt}\left(\frac{v_\alpha}{\left(1 - \frac{v^2}{c^2}\right)^{1/2}}\right) \end{aligned}\right\} \tag{6}$$

Substitution of relations (6) in the second term of Eq. 3 and addition of the same to (5) yields

$$\frac{\partial f}{\partial t} + \vec{v}\cdot\vec{\nabla}_r f + \frac{\left(1 - \frac{v^2}{c^2}\right)^{3/2}}{m_0} \sum_{\alpha=1}^{3} \frac{\partial f}{\partial v_\alpha}\frac{d}{dt}\left(\frac{m_0 v_\alpha}{\left(1 - \frac{v^2}{c^2}\right)^{1/2}}\right) = 0 \tag{7}$$

where m_0 is the rest mass of the particle.

Note that no velocity dependent forces are assumed to exist in the derivation of Eq. 7. It is clear that Eq. 7 of this paper, and Eq. 8 of (2) are identical but for a velocity dependent force term included by Clemmow and Willson. It is not necessary to show the invariance of Eq. 7 because the derivation presented here started from the proper frame and the Lorentz transformation is applied. However, if desired, a proof similar to that given in (2) may be stated. Noting that the ordinary force $\vec{F}$ is rate of change of momentum Eq. 7 may be written as

$$\frac{\partial f}{\partial t} + \vec{v} \cdot \vec{\nabla}_r f + \frac{\left(1 - \frac{v^2}{c^2}\right)^{3/2}}{m_0} \vec{F} \cdot \vec{\nabla}_v f = 0 \tag{8}$$

where

$$\vec{F} = \frac{d}{dt}\left(\frac{m_0 \vec{v}}{\left(1 - \frac{v^2}{c^2}\right)^{1/2}}\right).$$

In the problems of interest, $\vec{F}$ is the Lorentz force and is given by $e(\vec{E} + \vec{v} \times \vec{B})$, where e is the magnitude of the charge, $\vec{E}$ and $\vec{B}$ being the electric and the magnetic field intensities, respectively.

Note that in Eq. 8 the motion of positive ions is neglected. However, to take account of the same one can introduce another distribution function $g(\vec{x}, \vec{v}, t)$ and write an equation similar to Eq. 8 above. If one assumes, however, that the positive ions are uniformly smeared out, one need not consider $g(\vec{x}, \vec{v}, t)$.

Although there are many applications of Eq. 8, we will consider below the problem of a drifting electron beam. One comes across such a problem in the discussion of microwave devices like the klystron.

III. FORMULATION AND SOLUTION OF THE PROBLEM OF DRIFTING ELECTRON BEAM

To simplify the mathematics only one dimensional electron streams are considered here. Imagine a parallel plane (diode type) region across which is impressed (or exists) a small alternating (a.c.) electric field. The distribution function then satisfies the Boltzmann equation in one dimensional form

$$\frac{\partial f}{\partial t} + v \frac{\partial f}{\partial x} - \frac{\left(1 - \frac{v^2}{c^2}\right)^{3/2}}{m_0} eE \frac{\partial f}{\partial v} = 0 \tag{9}$$

where e is the electronic charge, x is the space coordinate normal to the planes, v is the x-directed velocity and E is the electric field. Let,

$$\left.\begin{aligned} E &= E_1 e^{i\omega t}, \\ f &= f_0(x, v) + f_1(x, v) e^{i\omega t} \end{aligned}\right\}, \tag{10}$$

so that $f_1 \ll f_0$ (f_0 the equilibrium distribution) and E_1 is assumed to be small. Substituting of relations (10) in Eq. 9 and linearizing (that is, neglecting the products of small quantities), one obtains Eq. 11 for the a.c. components of f_1.

$$i\omega f_1 + v\frac{\partial f_1}{\partial x} = \frac{\left(1 - \frac{v^2}{c^2}\right)^{3/2}}{m_0} eE_1 \frac{\partial f_0}{\partial v}. \tag{11}$$

From Maxwell's equations, one has

$$\frac{\partial E_1}{\partial x} = -\frac{1}{\epsilon_0}\int_{-\infty}^{\infty} f_1 dv. \tag{12}$$

Integrating Eq. 11 one obtains

$$f_1(x, v) = f_1(0, v)e^{-i\omega x/v} + \int_0^x \frac{\left(1 - \frac{v^2}{c^2}\right)^{3/2}}{m_0} v^{-1}E_1(x')\frac{\partial f_0}{\partial v}\exp\left[\frac{-i\omega(x - x')}{v}\right]dx'. \tag{13}$$

$f_1(x, v)$ can be determined from the integral equation (13) with a knowledge of $v_1(x)$, $f_1(0, v)$, $f_0(v)$ and $E_1(x)$. The only boundary condition needed is the value of $f_1(0, v)$. One can also obtain an integral equation for the a-c. current component as follows:

$$i_1(x) = \int_0^{\infty} v f_1(x, v)dv. \tag{14}$$

In a one dimensional problem, with no sources inside the plasma (or electron stream), from Maxwell's equations one obtains

$$E_1 = \frac{i_1}{i\omega\epsilon_0}. \tag{15}$$

Use of Eqs. 13, 14 and 15 yields the following integral equation for the a-c. component of the current,

$$i_1(x) = \int_0^{\infty} v f_1(0, v)\exp\left(\frac{-i\omega x}{v}\right)dv + \frac{e}{m_0 i\omega\epsilon_0}\int_0^x i_1(x')dx'\int_0^{\infty}\frac{\partial f_0}{\partial v}\left(1 - \frac{v^2}{c^2}\right)^{3/2} \times \exp\left[\frac{-i\omega(x - x')}{v}\right]dv. \tag{16}$$

Drifting Electron Beam with Small-Amplitude Velocity-Modulation

Although one can use the relativistic Maxwellian velocity distribution (7) for $f_0(v)$ in the solution of the present problem we will not do so here. Instead for simplicity in mathematics which can be justified (8), a rectangular distribution for $f_0(v)$ will be assumed here.

For the computation of $f_1(0, v)$, the model chosen is as follows. The stream is allowed to pass through two parallel grids across which there appears a voltage $V_1 e^{i\omega t}$. It is assumed that no transit time effects occur and the grids are perfectly permeable, $V_1 \ll V_0$, V_0 being the voltage corresponding to the initial velocity of the electron stream.

To the left of the gap, we have

$$f_0(v) = \frac{\rho_0}{w}[S(v - v_0) - S(v - v_0 - w)] \tag{17}$$

where ρ_0 is the d-c. charge density at the entrance plane, w the width of the distribution function, and

$$S(v - v_0) = \begin{cases} 1, & \text{for} \quad v \geqslant v_0 \\ 0, & \text{otherwise} \end{cases}$$

(v_0 is the smallest velocity of an electron in the drift region defined by the d-c. potential.)

Also,

$$v_0 = \frac{c[2\tau_0(1 + \tau_0/2)]^{1/2}}{(1 + \tau_0)}, \qquad \text{and} \qquad \tau_0 = \frac{eV_0}{m_0c^2}.$$

One can also write Eq. 17 in terms of voltages corresponding to the velocities as

$$f_0(v) = \frac{\rho_0}{w}[S(V - V_0) - S(V - V_0^{(w)})] \tag{17a}$$

where V_0 and $V_0^{(w)}$ are the potentials (voltages) corresponding to velocities v_0 and $(v_0 + w)$, respectively, and S is a step function already defined.

To the right of the gap, we have

$$f_0 + f_1 e^{i\omega t} = \frac{\rho_0}{w}\{S(V - V_0 - V_1 e^{i\omega t}) - S(V - V_0^{(w)} - V_1 e^{i\omega t})\}. \tag{18}$$

Expanding the right hand side of Eq. 18 in a Taylor series and retaining only terms of first order in V_1 $\left(\tau_1 = \dfrac{eV_1}{m_0c^2}\right)$, one obtains

$$f_1(0, v) \simeq -\frac{e}{m_0} V_1 \frac{\rho_0}{w}\left\{g(\tau_0)\frac{\delta(v - v_0)}{v_0} - g(\tau_0{}^{(w)})\frac{\delta(v - v_0 - w)}{(v_0 + w)}\right\} \quad (19)$$

where $g(\tau) = [1 + \tau(1 + \tau/2)](1 + \tau)^{-3}$ and δ denotes a Dirac délta function.

If it is further assumed that $g(\tau_0{}^{(w)}) \approx g(\tau_0)$ (that is, a narrow distribution is treated), then Eq. 19 further simplifies to Eq. 20 which will be employed in the ensuing calculations.

$$f_1(0, v) \simeq -\frac{e}{m_0} V_1 \frac{\rho_0}{w} g(\tau_0)\left\{\frac{\delta(v - v_0)}{v_0} - \frac{\delta(v - v_0 - w)}{(v_0 + w)}\right\}. \quad (20)$$

To obtain the a-c. component of current in a drifting electron beam which was initially velocity modulated, one has to solve Eq. 16 subject to the boundary condition given by Eq. 20. Equation 16 can be solved by Laplace transforms, and to do this introduce the Laplace transform defined by

$$i_1(p) = \int_0^\infty i_1(p)e^{-px}dx, \qquad \text{real part of } p > 0. \quad (21)$$

Taking the Laplace transform of Eq. 16 and using the Faltung theorem, one obtains

$$i_1(p) \simeq \left[\int_0^\infty \frac{v f_1(0, v)}{p + \frac{i\omega}{v}} dv\right] \times \left[1 - \frac{e\left(1 - \frac{v_0^2}{c^2}\right)^{3/2}}{m_0 i\omega\epsilon_0} \cdot \int_0^\infty \frac{\partial f_0}{\partial v} \cdot \frac{1}{\left(p + \frac{i\omega}{v}\right)} dv\right]^{-1}. \quad (22)$$

It is obvious that $i_1(x)$ can be obtained from

$$i_1(x) = \frac{1}{2\pi i}\int_{\beta - i\infty}^{\beta + i\infty} i_1(p)e^{px}dp. \quad (23)$$

We also have

$$\frac{\partial f_0}{\partial v} = \frac{\rho_0}{w}\{\delta(v - v_0) - \delta(v - v_0 - w)\}. \quad (24)$$

Substitution of Eqs. 20 and 24 in Eq. 22 yields, after integration,

$$i_1(p) \simeq \frac{i\omega V_1 \rho_0 e g(\tau_0)}{m_0 v_0^2 \left(1+\frac{w}{v_0}\right)} \left[\left(p+\frac{i\omega}{v_0}\right)\left(p+\frac{i\omega}{(v_0+w)}\right) + \frac{\omega_P^2}{v_0^2\left(1+\frac{w}{v_0}\right)} \right]^{-1} \quad (25)$$

where

$$\omega_P^2 = \frac{e\rho_0}{m_0\epsilon_0}\left(1-\frac{v_0^2}{c^2}\right)^{3/2}. \quad (26)$$

It is also assumed that $\left(1-\frac{v_0^2}{c^2}\right)^{3/2} \approx \left(1-\frac{(v_0+w)^2}{c^2}\right)^{3/2}$ or, $\frac{w}{v_0} \ll 1$.

Finding the poles of $i_1(p)$, and neglecting terms of the order greater than one in w/v_0, after a little manipulation one finds that

$$i_1(x) \simeq \frac{ieV_1\rho_0}{m_0 v_0} g(\tau_0)\left(1-\frac{w}{v_0}\right)\left(\frac{\omega}{\omega_P}\right) \times \left\{\sin\left(1-\frac{w}{v_0}\right)\frac{\omega_P x}{v_0}\right\} \exp\left[-i\omega\left(1-\frac{w}{v_0}\right)\frac{x}{v_0}\right]. \quad (27)$$

The non-relativistic analogue of Eq. 27 has been obtained previously in (8).

Although it is already known, it is not out of place to draw the attention of the reader to two facts:

(*a*) The plasma frequency that enters in is as defined by Eq. 26—where the so-called longitudinal mass of the electron is employed, and

(*b*) The presence of $g(\tau_0)$ (<1 in the relativistic case) indicates that it is much harder to modulate a relativistic electron beam than a non-relativistic one where $g(\tau_0)$ is essentially unity.

Finally it may be added that there are many other problems where the relativistic Boltzmann equation may be employed.

REFERENCES

(1) S. CHAPMAN AND T. G. COWLING, "The Mathematical Theory of Non-uniform Gases," second edition, New York, Cambridge University Press, 1952.

(2) P. C. CLEMMOW AND A. J. WILLSON, *Proc. Cambr. Phil. Soc.*, Vol. 53, p. 222 (1957).

(3) C. FLAMMER, unpublished work (has obtained among other things a dispersion relation similar to that given by Clemmow and Willson in *Proc. Roy. Soc. A*, Vol. 237, p. 117 (1956)).

(4) S. T. BELIAEV AND G. I. BUDKER, *Soviet Physics Doklady*, Vol. 1, p. 218, (1956).

(5) S. BLUDMAN AND K. M. WATSON, work to be published (the problems of stability of relativistic beams etc., have been treated).

(6) See for instance, L. D. LANDAU AND E. LIFSHITZ, "The Classical Theory of Fields," Reading (Mass.), Addison-Wesley Publishing Co., 1949.

(7) See for instance, J. L. SYNGE, "The Relativistic Gas," Amsterdam, North Holland Publishing Co., (1957).

(8) S. V. YADAVALLI, *Qu. App. Math.*, Vol. XII, p. 105 (1954); (a justification for the rectangular distribution based on the Tschebyscheff in equality was given which is applicable even in non-relativistic cases).

THEORY OF MAJORITY DECISION ELEMENTS

BY

SABURO MUROGA,[1] IWAO TODA AND SATORU TAKASU[2]

SUMMARY

A decision-making organization whose output is either 1 or 0 according to which of the numbers, 1 or 0, is predominant in its inputs, is defined as the majority decision element. A Boolean function represented by such an element is called a majority decision function. Elements of this sort play a significant role in the computer field. The present paper deals with various problems concerned with a single element of this sort: that is, in Section 3, the algebraic properties, and a necessary condition for realizability of a given function; in Section 4, the types of functions realizable by a single element; in Section 5, the determination of the structure of the element for a given function by means of linear programming; and in Section 6, properties of the functions of a small number of variables realizable by a single element.

1. INTRODUCTION

The neurons in living organisms are believed to work on a majority principle and their behavior can be analyzed by propositional logic (1).[3] It has been shown that the majority decision principle reduces failures of neural signal transmission (2).

Recently in electronic computers, logic elements based essentially on the majority decision principle, like magnetic cores (3) and parametrons (4), have been used. Even circuits with vacuum tubes or transistors could be built on this principle and actually the principle had been used in the Kirchhoff adder. One of the authors has showed (5) that the idea of coupling numbers, that is, relative coupling amplitudes of inputs, makes even a single element based on the majority decision principle represent a fairly complex function and that a network for a given function can be realized with fewer elements of this sort and with a higher speed. As the switching speed of the physically realizable elements is limited, this realizability of a complex function by a single element may be of some importance for speeding up computers. Also interesting relations were shown between the probabilistic behavior of the element with its representable functions, and an error-reducing scheme of a network of such elements (6).

---

[1] IBM Research Center, Yorktown Heights, N. Y.; formerly with the Electrical Communication Laboratory, Nippon Telegraph and Telephone Public Corp., Tokyo, Japan.

[2] The Electrical Communication Laboratory, Nippon Telegraph and Telephone Public Corp., Tokyo, Japan.

[3] The boldface numbers in parentheses refer to the references appended to this paper.

2. MATHEMATICAL REPRESENTATION OF THE MAJORITY DECISION ELEMENT

Definition of the Majority Decision Element

In the present paper, two states of firing and nonfiring of the neuron, and two different voltage values of a logic element in a computer are expressed by the binary values, 1 and 0. Of course there are other expressions like $+1$ and -1. A decision-making organization whose output is either 1 or 0 according to which of the numbers, 1 or 0, is predominant in its inputs, will be henceforth called the majority decision element.

Generalizing this in this paper, a value weighted on the voting effect of each input value will be called a "coupling number" of the input. The majority decision will depend on this. Let coupling numbers of n variable binary inputs $x_1, x_2, \cdots, x_n$ be $w_1, w_2, \cdots, w_n$, respectively, and let w_c be a coupling number of a constant input that takes permanently the value 1 or 0. Then an output value of the element is specified as 0 or 1 according to which of a sum of coupling numbers of inputs having 0's and that of inputs having 1's is larger. However it may be convenient to use the concept of a threshold T of the element as follows and to check whether a sum of coupling numbers of "variable" inputs having 1's reaches this threshold value or not.

w_i $(i = 1, 2, \cdots, n)$ in this paper is assumed to be a positive real number. When coupling of an input to an element is realized physically by wiring on a transformer as in the parametron, it is allowed only to be a positive or a negative integer (**4**). In a circuit with the Esaki diodes to which three-cycle excitation is applied (**7**, **8**), coupling is realized with a resistor and it is a positive real number. However as a negative value of w_i can be reduced to a case of a positive w_i by taking a negation of the input as a new variable, negativeness of w_i is not essential as far as a single element of the majority decision principle is dealt with and will be henceforth eliminated. Thus, in this paper w_i is limited to a positive real number and in most cases a positive integral value of w_i will be considered.

Definition 1: Let coupling numbers of variable inputs $x_1, x_2, \cdots, x_n$ be $w_1, w_2, \cdots, w_n$, and a coupling number of a constant input be w_c. The sum of these will be called the total coupling number of inputs:

$$K = w_1 + w_2 + \cdots + w_n + w_c. \tag{1}$$

Definition 2: A majority decision element is defined as an element whose output value is

$$1, \text{ if } \sum_{i=1}^{n} w_i x_i \geqq T \tag{2a}$$

$$0, \text{ if } \sum_{i=1}^{n} w_i x_i < T, \tag{2b}$$

for n variable inputs $x_1, x_2, \cdots, x_n$. Coupling numbers $w_1, w_2, \cdots, w_n$ and T are assumed to exist and to be positive real numbers.

When we can find a positive number ϵ such that $\sum_{i=1}^{n} w_i x_i \leqq T - \epsilon$ always holds as (2b), (2a) and (2b) may be reduced to (4a) and (4b), given later, by dividing (2a) and (2b) by ϵ and replacing w_i/ϵ and T/ϵ by new w_i and T. Therefore (4a) and (4b) can be used henceforth instead of (2a) and (2b). Note that the x_i's in (2a) and (2b) are no longer variables but take fixed values of 1 or 0.

The behavior of a majority decision element can be expressed by a Boolean function. For each subscript set $\alpha, \beta, \cdots, \gamma$ such that $w_\alpha + w_\beta + \cdots + w_\gamma \geqslant T$, construct a product $x_\alpha x_\beta \cdots x_\gamma$. The logical sum of all of such products represents the Boolean function that expresses the majority decision element in Definition 2. This will be called the *majority decision function*. Note that any product that includes another product as a factor can be deleted from the function, because the latter always takes the value of 1 when the former is 1. Thus we have a Boolean polynomial without negation of the variables. Henceforth throughout this paper we assume that *a majority decision function is in this irredundant form, that is, consists of only prime implicants.*

The structure of a majority decision element is completely specified by its threshold T and the coupling numbers of variable inputs, $w_1, w_2, \cdots, w_n$. The structure will henceforth be denoted simply as $[w_1, w_2, \cdots, w_n : T]$. Alternatively, specification of w_e instead of T also can specify the structure. It is easy to get a Boolean polynomial or the inequalities in Definition 2 from this structure, but in general, determination of the structure from a given Boolean polynomial is not easy; it is a major theme of the present paper. The threshold T may be a number taken such that $S_M < T \leqq S_N$. Here S_M is the maximum of sums of coupling numbers of the inputs, each of which does not reach the threshold with values of one of these inputs, and S_N is the minimum of those, each of which exceeds the threshold with values of one of these inputs.

However, when all of $w_1, w_2, \cdots, w_n$ and w_e are positive integers, and if the total coupling number of the inputs, K, is an odd integer, the threshold T may be defined as a positive integer as follows:

$$T = \tfrac{1}{2}[(w_1 + w_2 + \cdots + w_n) + 1 \pm w_e], \tag{3}$$

where $-$ and $+$ correspond to 1 and 0 of a constant input, respectively. Then $\sum_{i=1}^{n} w_i x_i$ takes only positive integer values, so Eqs. 2a and 2b may be replaced by the following: An output value of the element is

$$1, \text{ if } \sum_{i=1}^{n} w_i x_i - T \geqq 0$$

$$0, \text{ if } -\sum_{i=1}^{n} w_i x_i + T \geqq 1.$$

This paper will be based on this form of equations. Let concrete values of x_i's be expressed as x_{ij} and then the above may be written as follows: An output of the element is

$$1, \text{ if } \sum_{i=1}^{n} w_i x_{ij} - T \geqq 0 \qquad (j = 1, 2, \cdots, r) \tag{4a}$$

$$0, \text{ if } -\sum_{i=1}^{n} w_i x_{ij} + T \geqq 1 \qquad (j = r+1, \cdots, 2^n). \tag{4b}$$

Relation Between the Physical Elements Based on Majority Decision Principle and Their Theoretical Models

In the elements of Definition 2, variables take values either of 1 or 0 and outputs are decided by the majority of these binary values. Physically realized elements, however, like parametrons, neurons and Esaki-diode circuits, may not be interpreted to work concretely on values of 1 and 0. It may be physically more accurate to state that parametrons work on the majority of $+1$ and -1 (4). In the case of neurons, it may be natural to regard 1 and 0 for excitatory inputs, and 0 and -1 for inhibitory inputs, as the contribution of the majority decision (2).

In the case of a neuron with both excitatory inputs and inhibitory inputs, an input may appeapto take ternary values of 1, 0, and -1, but actually each of the inputs takes binary values, that is, 1 and 0, respectively for 1 and 0 of an excitatory input and -1 and 0, respectively, for 1 and 0 of an inhibitory input. Then a theoretical model for the neuron where all inputs take the same binary values of 1 and 0 can be obtained as follows. When the number of inhibitory inputs is m, replace the threshold value T by $T + m$ and take 0 and 1 instead of -1 and 0 of each inhibitory input, respectively, as the contribution to the majority decision.

Situations similar to these appear in various circuits with the Esaki diodes.

However, consideration of the binary values as being 1 and 0 is sufficient to treat theoretically the above physical devices. As shown in the following, it is easy to convert between values of input variables x_i's in the theoretical model and those of the physical quantity ξ_i's that participate in majority decision of the output of the physical device.

Suppose that ξ_i assumes physical values either of a_i or b_i correspond-

ing to 1 or 0 of x_i. Then we have the following conversion formula

$$\xi_i = a_i x_i + b_i(1 - x_i). \tag{5}$$

In the physical device based on the majority decision principle, its output assumes the value

$$1, \text{ if } \sum w_i \xi_i \geqq T$$
$$0, \text{ if } \sum w_i \xi_i \leqq T - 1.$$

Substitution of Eq. 5 gives the relation that an output of the element assumes the value

$$1, \text{ if } \sum w_i(a_i - b_i)x_i \geqq T - \sum b_i$$
$$0, \text{ if } \sum w_i(a_i - b_i)x_i \leqq T - \sum b_i - 1.$$

This represents a theoretical model with a threshold $T - \sum b_i$ and with coupling numbers $w_i(a_i - b_i)$ for x_i's, corresponding to the above device. Conversely, the structure of the physical device can be specified from that of the theoretical model if a_i and b_i of the device are known.

Consequently, the theoretical model of Definition 2 is sufficient for discussions of the logical behavior of any physical devices based on the majority decision principle.

3. SOME PROPERTIES OF THE MAJORITY DECISION FUNCTION AND A NECESSARY CONDITION FOR ITS REALIZABILITY BY A SINGLE ELEMENT

Real numbers named as coupling numbers are attached to inputs of the majority decision element and ordinal relations between them specify the behavior of the element. Here we shall discuss how a relation between coupling numbers and a threshold is reflected in a property of the majority decision function itself.

An inclusion relationship between Boolean functions may be defined as follows: Consider two functions $f(x_1, \cdots, x_n)$ and $g(x_1, \cdots, x_n)$; if any set of $x_1, x_2, \cdots, x_n$ satisfying $f(x_1, \cdots, x_n) = 1$ satisfies also $g(x_1, \cdots, x_n) = 1$, but the converse does not necessarily hold, we write

$$f(x_1, \cdots, x_n) \subseteq g(x_1, \cdots, x_n). \tag{6}$$

In addition, if Eq. 6 holds and, further, there exists a set of $x_1, \cdots, x_n$ satisfying simultaneously $f(x_1, \cdots, x_n) = 1$ and $g(x_1, \cdots, x_n) = 0$, we write

$$f(x_1, \cdots, x_n) \subset g(x_1, \cdots, x_n). \tag{7}$$

A function $f(x_1, \cdots, x_n)$, where for any $i (i = 1, 2, \cdots, n)$,

$$f(x_1, \cdots, \overset{i}{\breve{1}}, \cdots, x_n) \supseteq f(x_1, \cdots, \overset{i}{\breve{0}}, \cdots, x_n) \tag{8}$$

holds, is called a positive function. A positive function can be expressed by a polynomial of $x_1, \cdots, x_n$ without negations of these variables and vice versa. A majority decision function is a positive function but the converse is not true (5). The positive functions defined here are the same as E. N. Gilbert's switching functions that are expressed without negations of variables (9).

Ordinal Relations Among Input Variables

A structure of the majority decision element specifies the function that the element represents. However, the structure of an element that is to represent a given majority decision function is not unique. Consider the following example. The majority decision function $x_1 + x_2$ is represented by the element with the structure $w_1 = w_2 = 1$ and $T = 1$ and also by $w_1 = 2, w_2 = 3$ and $T = 2$. However, it may be important to see as in Theorem 2 that when a majority decision function is given, at least ordinal relations among coupling numbers can be uniquely determined with respect to any one of the elements that are to represent the given function. The following definition will be useful.

Definition 3: Consider a positive function $f(x_1, x_2, \cdots, x_n)$; a relation

$$f(x_1, \cdots, \overset{i}{\underset{\smile}{1}}, \cdots, \overset{j}{\underset{\smile}{0}}, \cdots, x_n) \supseteq f(x_1, \cdots, \overset{i}{\underset{\smile}{0}}, \cdots, \overset{j}{\underset{\smile}{1}}, \cdots, x_n) \tag{9}$$

for any set of i and $j(\neq i)$, is denoted as

$$x_i \gtrsim x_j. \tag{10}$$

Verbally it may be stated as "the input variable x_i is not smaller than x_j." In particular, when an equality is not included in (9), it is denoted as $x_i > x_j$ and when only equality is included there, $x_i \sim x_j$. For completeness of definition, $x_i \sim x_i$ is defined.

This binary relation has the following properties:

Theorem 1: The binary relation $\gtrsim$ on the positive functions is a partial ordering.

Proof: Obviously this binary relation satisfies the reflexive and anti-symmetric laws from the definition. Assume $x_1 \gtrsim x_2$ and $x_2 \gtrsim x_3$, that is,

$$\begin{aligned} &f(1, 0, x_3, \cdots, x_n) \supseteq f(0, 1, x_3, \cdots, x_n) \\ &f(x_1, 1, 0, x_4, \cdots, x_n) \supseteq f(x_1, 0, 1, x_4, \cdots, x_n). \end{aligned} \tag{11}$$

From a property of the positive functions

$$\begin{aligned} f(1, x_2, 0, x_4, \cdots) &= f(1, 1, 0, x_4, \cdots)x_2 + f(1, 0, 0, x_4, \cdots) \\ f(0, x_2, 1, x_4, \cdots) &= f(0, 1, 1, x_4, \cdots)x_2 + f(0, 0, 1, x_4, \cdots) \end{aligned} \tag{12}$$

hold. Inserting $x_2 = 1$ and $x_1 = 1$ in the two inequalities of (11), respectively, we have $f(1, 1, 0, x_4, \cdots, x_n) \supseteq f(1, 0, 1, x_4, \cdots, x_n) \supseteq f(0, 1, 1, x_4, \cdots, x_n)$, and thus the first term in the first expansion in (12) is not smaller than that in the second expansion. Similarly, the second terms are compared by inserting $x_2 = 0$ and $x_1 = 0$ in (11). Consequently, as the terms in the first expansion of (12) were shown to be not smaller than the corresponding terms in the second expansion, we have $f(1, x_2, 0, x_4, \cdots, x_n) \supseteq f(0, x_2, 1, x_4, \cdots, x_n)$, that is, $x_1 \gtrsim x_2$ and the transitive law is satisfied.

When we proceed into the majority decision functions from the positive functions, this relation will prove to be a total ordering.

Theorem 2: The binary relation $\gtrsim$ for a majority decision function $f(x_1, x_2, \cdots, x_n)$ is a total ordering. w_i's that have an ordinal relation among them as real numbers are homomorphic with the input variables x_i's defined with the binary relations $\gtrsim$. That is, if $w_i > w_j$, then

$$x_i \gtrsim x_j \tag{13}$$

and if $w_i = w_j$, then

$$x_i \sim x_j. \tag{14}$$

Proof: As the relations were shown to be a partial ordering in Theorem 1, if either $x_i \gtrsim x_j$ or $x_i \lesssim x_j$ for any set of x_i and x_j holds, the relations are total orderings. If the majority decision function $f(x_i, x_2, \cdots, x_n)$ is realized with a single element of a structure $[w_1, w_2, \cdots, w_n; T]$, one of the following three relations must hold between coupling numbers w_i and w_j of any two variables:

(i) When $w_i > w_j$, if for a set of values of variables excluding x_i and x_j

$$\sum_{k \neq i, j} w_k x_k \geqq T - w_i, \tag{15}$$

holds, $f(x_1, \cdots, \overset{i}{\check{1}}, \cdots, \overset{j}{\check{0}}, \cdots, x_n)$ takes value 1. And if

$$\sum_{k \neq i, j} w_k x_k \geqq T - w_j \tag{16}$$

holds, $f(x_1, \cdots, \overset{i}{\check{0}}, \cdots, \overset{j}{\check{1}}, \cdots, x_n)$ takes value 1. Since $T - w_i$

$< T - w_j$ holds from the assumption, a set of variable values (excluding x_i and x_j) which satisfies (16) satisfies (15) also and therefore

$$f(x_1, \cdots, \overset{i}{\underset{\smile}{1}}, \cdots, \overset{j}{\underset{\smile}{0}}, \cdots, x_n) \supseteq f(x_1, \cdots, \overset{i}{\underset{\smile}{0}}, \cdots, \overset{j}{\underset{\smile}{1}}, \cdots, x_n) \quad (17)$$

holds. Consequently we have $x_i \gtrsim x_j$.

(ii) When $w_i = w_j$, we can conclude by a discussion similar to that of case (i) that only an equality holds in (17) and therefore $x_i \sim x_j$.

(iii) When $w_i < w_j$, we again, as with case (i), have $x_i \lesssim x_j$.

Consequently, since for any x_i and x_j a binary relation $\lesssim$ exists, the homomorphism of x_i defined in this way with w_i's as real numbers has been proved.

From this theorem, input variables x_i's can be arranged in order by means of the binary relation $\gtrsim$. Even though the structure that is actually required to realize a given function is not known, ordinal relations among coupling numbers to be realized can be seen from the binary relations peculiar to the function. Of course if $x_i > x_j$, then $w_i > w_j$ holds. For example, for the function $f(x_1, x_2, x_3) = x_1 + x_2 \cdot x_3$, as $f(1, 0, x_3) = 1 \supset f(0, 1, x_2) = x_3$ and $f(x_1, 0, 1) = x_1 = f(x_1, 1, 0)$ hold, we have $x_1 > x_2 \sim x_3$. This function is actually realized with an element of structure $[2, 1, 1; 2]$ where there are the relations $w_1 > w_2 = w_3$.

The following corollary is convenient to check whether a given positive function is actually a majority decision function (17).

Corollary 1: A majority decision function $f(x_1, \cdots, x_n)$ can be written in the following form:

$$f(x_1, \cdots, x_n) = M(x_1, \cdots, \overset{i}{\underset{\smile}{\ }}, \cdots, x_n)x_i + N(x_1, \cdots, \overset{i}{\underset{\smile}{\ }}, \cdots, x_n) \quad (18)$$

where x_i is any one of the largest variables compared in the binary relation $\gtrsim$. Here a relation

$$M \supset N \quad (19)$$

holds. Both $M(x_1, \cdots, \overset{i}{\underset{\smile}{\ }}, \cdots, x_n)$ and $N(x_1, \cdots, \overset{i}{\underset{\smile}{\ }}, \cdots, x_n)$ represent the majority decision functions of $(n-1)$ variables excluding x_i. These can be further expanded as follows:

$$M(x_1, \cdots, \overset{i}{\smile}, \cdots, x_n) = m_1(x_1, \cdots, \overset{i}{\smile}, \cdots, \overset{j}{\smile}, \cdots, x_n)x_j + m_2(x_1, \cdots, \overset{i}{\smile}, \cdots, \overset{j}{\smile}, \cdots, x_n)$$
$$N(x_1, \cdots, \overset{i}{\smile}, \cdots, x_n) = n_1(x_1, \cdots, \overset{i}{\smile}, \cdots, \overset{j}{\smile}, \cdots, x_n)x_j + n_2(x_1, \cdots, \overset{i}{\smile}, \cdots, \overset{j}{\smile}, \cdots, x_n), \tag{20}$$

with respect to a certain $x_j (j \neq i)$ that exists always and the relations

$$m_1 \supset m_2 \supseteq n_1 \supset n_2 \tag{21}$$

hold. $m_1(x_1, \cdots, \overset{i}{\smile}, \cdots, \overset{j}{\smile}, \cdots, x_n)$ etc. mean functions of $(n-2)$ variables excluding x_i and x_j.

If we assume $w_i \geq w_j \geq \cdots$, the proof is easily obtained since f consists of prime implicants.

Binary Relations among Input Vectors

In the previous section the binary relations are defined by giving concrete values of 1 and 0 to only two of n variables. Here, generalizing this, 1 and 0 will be given to all n variables. While in the previous section coupling numbers of input variables are compared pairwise, the treatment in this section implies comparison of sums of coupling numbers of input variables having value 1. When binary values are concretely assigned to all n variables, those sums of coupling numbers of variables having value 1, which exceed the threshold, specify the majority decision functions for the element. Though the ordinal relations among these sums, for various sets of values of n variables, can be easily determined if a structure of the element is known, they can be specified to some extent from properties of the given function even if it is not known, as shown in Section 3 of Theorem 3. However as shown in Section 2 of Theorem 3, this relation is not a true ordering, but a "pseudo ordering."

Concrete values of input variables x_i's will be assigned by ξ_i's and a particular set of these values by a vector expression $\Xi = (\xi_1, \cdots, \xi_n)$. This will be called an "input vector." The j^{th} vector in these 2^n input vectors will be written as:

$$\Xi_j = (\xi_1^{(j)}, \xi_2^{(j)}, \cdots, \xi_n^{(j)}) \tag{22}$$

where $j = 0, 1, \cdots, 2^n - 1$ and $\xi_i^{(j)} = 0$ or 1 $(i = 1, 2, \cdots, n)$.

Definition 4: Let Ξ_1 and Ξ_2 be two input vectors, where m out of n components of $\xi_i^{(1)}$ and $\xi_i^{(2)}$ $(i = 1, 2, \cdots, n)$ are of the same values. For example,

$$\begin{aligned} &\xi_i^{(1)} = \xi_i^{(2)} \qquad (i = 1, 2, \cdots, m) \\ &\xi_i^{(1)} \neq \xi_i^{(2)} \qquad (i = m+1, \cdots, n) \end{aligned} \tag{23}$$

(in general i's for these equalities and inequalities are not necessarily consecutive). If

$$f(x_1, \cdots, x_m, \xi_{m+1}^{(1)}, \cdots, \xi_n^{(1)}) \supseteq f(x_1, \cdots, x_m, \xi_{m+1}^{(2)}, \cdots, \xi_n^{(2)}) \tag{24}$$

holds for two functions of $x_1, \cdots, x_m$ out of the n variables, this fact will be denoted by

$$\Xi_1 \overset{f}{\geq} \Xi_2, \tag{25}$$

where equality and inequality in (24) and (25) correspond to each other.

Though this binary relation may be reduced to $\gtrsim$ (see Definition 3) if only two components differ in $\xi_i^{(1)}$ and $\xi_i^{(2)}$, a different notation $\overset{f}{\geq}$ is used here to stress a generalized idea. Such variables as $x_1, \cdots, x_m$ in this definition may be called free variables and the rest constrained variables. For example, when $f(x_1, x_2, x_3) = x_1x_2 + x_2x_3 + x_3x_1$ is given, for $\Xi_1 = (0, 1, 0)$ and $\Xi_2 = (1, 1, 1)$, x_2 is a free variable and x_1 and x_3 constrained variables. As $f(0, x_2, 0) = 0$ and $f(1, x_2, 1) = 1$, we have

$$\Xi_1 \overset{f}{<} \Xi_2. \tag{26}$$

Theorem 3: If $f(x_1, \cdots, x_n)$ is a majority decision function, the following holds:

1. A binary relation is always determinable between any two of the 2^n input vectors.

2. If $\Xi_1 \overset{f}{<} \Xi_2$ and $\Xi_2 \overset{f}{<} \Xi_3$, then $\Xi_1 \overset{f}{\leq} \Xi_3$ (note that this is not a truly ordinal relation).

3. In the special case where a structure $[w_1, w_2, \cdots, w_n; T]$ of a majority decision element realizing this function is known: if $\sum_{i=1}^{n} w_i\xi_i^{(1)} > \sum_{i=1}^{n} w_i\xi_i^{(2)}$ then

$$\Xi_1 \overset{f}{\geq} \Xi_2 \tag{27a}$$

and if $\sum_{i=1}^{n} w_i\xi_i^{(1)} = \sum_{i=1}^{n} w_i\xi_i^{(2)}$ then

$$\Xi_1 \overset{f}{=} \Xi_2. \tag{27b}$$

Proof: Let us start from Section 3 above. In the comparison of Ξ_1 and Ξ_2, assume $x_1, \cdots, x_m$ to be free variables and the rest to be constrained variables. As a structure of the element is given, $\sum_{i=1}^{n} w_i\xi_i^{(j)}$ for $j = 1$ and 2 can be calculated. Since $\sum_{i=1}^{m} w_i\xi_i^{(1)} = \sum_{i=1}^{m} w_i\xi_i^{(2)}$ holds, if one of $>$, $=$ and $<$ exists between the remaining $\sum_{i=m+1}^{n} w_i\xi_i^{(1)}$ and $\sum_{i=m+1}^{n} w_i\xi_i^{(2)}$, then one of $\supseteq$, $=$ and $\subseteq$ holds, respectively, between $f(x_1, \cdots, x_m, \xi_{m+1}^{(1)}, \cdots, \xi_n^{(1)})$, and $f(x_1, \cdots, x_m, \xi_{m+1}^{(2)}, \cdots, \xi_n^{(2)})$, by a treatment similar to the proof of Theorem 2. Thus Section 3 of Theorem 3 is proved. Next, if it is a majority decision function, there should exist coupling numbers of an element to realize it, for which the property described above holds for any set of Ξ_1 and Ξ_2. Then always a binary relation $\overset{f}{\geq}$ can be established and Section 1 of Theorem 3 is proved. Last, for a structure of an element to realize $f(x_1, \cdots, x_n)$, if $\Xi_1 \overset{f}{<} \Xi_2$ and $\Xi_2 \overset{f}{<} \Xi_3$ hold, we have $\sum_{i=1}^{n} w_i\xi_i^{(1)} < \sum_{i=1}^{n} w_i\xi_i^{(2)} < \sum_{i=1}^{n} w_i\xi_i^{(3)}$ from Section 3 of Theorem 3. Thus $\Xi_1 \overset{f}{\leq} \Xi_3$ is shown, proving Section 2.

From this theorem, the structure of an element realizing a given majority decision function can be seen to some extent. For example, assume that $f(x_1, \cdots, x_4) = x_1x_2 + x_1x_3x_4$ is given and compare $\Xi_1 = (1, 1, 0, 0)$ and $\Xi_2 = (0, 1, 1, 1)$, where only x_2 is a free variable and the rest constrained variables. Then $f(1, x_2, 0, 0) = x_2$ and $f(0, x_2, 1, 1) = 0$. As $x_2 \supset 0$, $(1, 1, 0, 0) \overset{f}{>} (0, 1, 1, 1)$ holds. From Section 3 of Theorem 3, if this function can be realized with a majority decision element, a relation $w_1 > w_3 + w_4$ must hold among its coupling numbers. Actually a structure of an element for this function is $[3, 2, 1, 1; 5]$ (16).

When there exists only one constrained variable, this theorem is merely a definition of the positive functions. When Theorems 2 and 3 are compared, their physical implications may be different, as the former compares individual coupling numbers pairwise but the latter compares various sums of coupling numbers. However, if regarded as a necessary condition for realizability of a given function by a single majority decision element, the latter is a stronger statement because the former always holds if the latter holds, but the converse does not necessarily hold. In Theorem 2 only a binary relation between any two variables is a necessary condition and its satisfaction of the ordering axiom is a necessary condition. Let us give an example of a function that does not satisfy Theorem 3 though it satisfies the condition of Theorem 2. For $f(x_1, \cdots, x_6) = \{(x_3 + x_4 + x_5 + x_6)x_2 + (x_4 + x_5 + x_6)x_3 + x_4x_5 + x_5x_6 + x_4x_6\}x_1 + x_2x_3x_4$, relations $x_1 > x_2 \sim x_3 \sim x_4 > x_5 \sim x_6$ can

be easily ascertained. But as $f(1, 0, 0, x_4, x_5, x_6) = x_4x_5 + x_5x_6 + x_4x_6$ and $f(0, 1, 1, x_4, x_5, x_6) = x_4$ hold, there does not exist an inclusion relationship between these two. Thus this function satisfies Theorem 2 but does not satisfy Theorem 3. Also a necessary condition obtained previously (5) compares actually individual coupling numbers and then is similar to Theorem 2. As shown later in the next section, there always exists an inclusion relation between a function satisfying Section 1 of Theorem 3 and its dual function. Finally, a necessary condition as the strongest among the ones obtained so far may be stated as follows: *A necessary condition* for realizability of a given function by a single majority decision element is Sections 1 and 2 of Theorem 3.

Duality of the Majority Decision Functions

When a function $f(x_1, \cdots, x_n)$ is realized by an element with the structure $[w_1, \cdots, w_n; T]$, its dual function $f^*(x_1, \cdots, x_n)$ is realized by an element with the structure $[w_1, \cdots, w_n; 1 - T + \sum_{i=1}^{n} w_i]$, by negating a value of a constant input (that is, if it is 1, convert it into 0 and *vice versa*) without change in coupling numbers of the constant input and variables (5). The element represents the same function since the element works on the majority of a number of 0's instead of that of 1's when all inputs and an output of the element are negated, that is, 1's and 0's are reversed. Now we will show that in any function satisfying condition 1 of Theorem 3 and further in the majority decision functions there always exists an inclusion relation between f and f^*.

Theorem 4: Consider a function $f(x_1, \cdots, x_n)$; if the binary relation defined in Definition 4 holds between any two of 2^n input vectors, either $f \supseteq f^*$ or $f \subseteq f^*$ holds.

Proof: A vector where all components of an input vector Ξ are negated will be denoted by $\bar{\Xi}$ (for example, $\bar{\Xi} = (0, 1, 1)$ for $\Xi = (1, 0, 0)$). Let A be the set of all input vectors which gives f the value of 1, and B be the set of all input vectors which gives $\bar{f}$ the value of 1, where $\bar{f}$ means negation of f. Also, let A^* and B^* be similar sets of input vectors for f^* and $\bar{f}^*$, respectively. Define $A' = \{\bar{\Xi} | \Xi \epsilon A\}$ and $B' = \{\bar{\Xi} | \Xi \epsilon B\}$ and then from the definition of the dual functions we have

$$A^* = B' \qquad \text{and} \qquad B^* = A'. \tag{28}$$

Now we will show that both

$$\begin{aligned} A \cap A' &\neq \phi \\ B \cap B' &\neq \phi \end{aligned} \tag{29}$$

do not hold simultaneously if the assumption of the theorem is satisfied, where ϕ means the null set. Assume that both hold simultaneously.

Then there should exist input vectors Ξ_1 and Ξ_2 such that $\Xi_1 \epsilon A$ and $\bar{\Xi}_1 \epsilon A$, $\Xi_2 \epsilon B$ and $\bar{\Xi}_2 \epsilon B$. As the free variables are obviously the same in the comparison of Ξ_1 and Ξ_2 as in the comparison of $\bar{\Xi}_1$ and $\bar{\Xi}_2$, decomposition of the components in these four vectors leads to $\Xi_1 = (X, \Phi)$, $\bar{\Xi}_1 = (\bar{X}, \bar{\Phi})$, $\Xi_2 = (X, \Psi)$ and $\bar{\Xi}_2 = (\bar{X}, \bar{\Psi})$, where X is a vector consisting of the free variables, Φ and Ψ are vectors consisting of the constrained variables and $\bar{X}$, $\bar{\Phi}$ and $\bar{\Psi}$ are their component-wise negated vectors. As all components corresponding to the constrained variables in Ξ_1 and Ξ_2 differ, we have $\Phi = \bar{\Psi}$. From the assumption, binary relations hold between Ξ_1 and Ξ_2 and between $\bar{\Xi}_1$ and $\bar{\Xi}_2$, and for vectors belonging to A and B we have always $f = 1$ and $f = 0$, respectively. Consequently, with X representing agreeing variables, there must hold

$$\begin{aligned} f(X, \Phi) &\supset f(X, \Psi) \\ f(X, \bar{\Phi}) &\supset f(X, \Psi) \end{aligned} \tag{30}$$

without equalities. However, as this contradicts the relation $\Phi = \bar{\Psi}$, at least one of assumed relations must be an equality. If

$$B \cap B' = \phi \tag{31}$$

is assumed, A and B divide 2^n input vectors into two parts and then we have

$$A \supseteq B'. \tag{32}$$

With (28) this becomes $A \supseteq A^*$. Then $f \supseteq f^*$ holds. Similarly, $f \subseteq f^*$ for $A \cap A' = \phi$.

For the self-dual functions, $A \cap A' = \phi$ and $B \cap B' = \phi$ in (29) hold simultaneously.

From this we have the following observation.

Corollary 2: Between a majority decision function f andi ts dual function f^*, either $f \supseteq f^*$ or $f \subseteq f^*$ always holds.

Variation of the Threshold

Here we shall discuss how the majority decision function represented by an element is affected by variation of its threshold, while fixing the coupling numbers of variable inputs. For simplicity, in this section all coupling numbers and the threshold are assumed as integers. A majority decision function for an element with a threshold t will be denoted as $f_t(x_1, \cdots, x_n)$.

Theorem 5: If $i \geq j$, $f_i(x_1, \cdots, x_n) \subseteq f_j(x_1, \cdots, x_n)$ holds.

Proof: Only when $x_1, \cdots, x_n$ satisfy $\sum_{k=1}^{n} w_k x_k \geq i$, f_i takes the value of 1. Then from the assumption, $\sum_{k=1}^{n} w_k x_k \geq i \geq j$ is satisfied and for these $x_1, \cdots, x_n$, $f_j = 1$ holds.

Theorem 6: The number of different functions that are realized by changing the threshold of an element but fixing the coupling numbers of variable inputs, is not greater than $2^n + 1$, including two functions which are permanently 1 and 0.

Proof: There are $2^n - 1$ different sums of coupling numbers, $w_1, w_2, \cdots, w_n; w_1 + w_2, \cdots; \cdots; w_1 + w_2 + \cdots + w_n$. These sums can be arranged according to their magnitudes and in general some of them may be equal. This greatest number of different sums can be actually obtained as follows. If we set $w_i = 2^{i-1}$, then the successive magnitudes each differ by 1 and then variation of the threshold within this range gives at most $2^n - 1$ different functions because a different T gives a different bisection of $2^n - 1$ sums, resulting in a different function. Here we have the function 0 for $T = 2^n$ and the function 1 for $T = 0$. In the case of $w_i = 2^{i-1}$, the forms of functions for a given T can be easily obtained.

For example, set $w_1 = 1$, $w_2 = 2$, and $w_3 = 4$. If we increase T one by one, we have functions $x_1 + x_2 + x_3$, $x_2 + x_3$, $x_1x_2 + x_3$, x_3, $x_1x_3 + x_2x_3$, x_2x_3 and $x_1x_2x_3$, accordingly. (Note that only the function $x_1x_2 + x_1x_3 + x_2x_3$ is excluded from all the majority decision functions of three variables.) However, in general any one of these functions can be realized with smaller coupling numbers.

Convolution of Structures of Two Majority Decision Elements

In this paper, convolution of two elements means making a majority decision element with coupling numbers $w_1 + w_1', \cdots, w_n + w_n'$, from the one element with $w_1, \cdots, w_n$ and the other element with $w_1', \cdots, w_n'$. In this section all coupling numbers and thresholds are assumed as integers.

Theorem 7: Let $m_t(x_1, \cdots, x_n)$ and $n_t(x_1, \cdots, x_n)$ be majority decision functions of two elements with two sets of coupling numbers, $w_1, \cdots, w_n$ and $w_1', \cdots, w_n'$, respectively, and threshold t. The majority decision function, represented by an element with a threshold T which is made by convolution of these two elements, can be written as follows:

$$f_T(x_1, \cdots, x_n) = \sum_{t=0}^{M} m_t(x_1, \cdots, x_n) n_{T-t}(x_1, \cdots, x_n) \tag{33}$$

where $M = \sum_{i=1}^{n} w_i$ and where, for $t \leqq 0$, m_t and n_t are regarded as the function which takes the value of 1 permanently.

Proof: The polynomial f_T is a sum of products, each of which consists of variables with values of 1 in an input vector such that

$\sum_{i=1}^{n}(w_i + w_i')x_1 \geqq T$ is satisfied. Let S be a set of such input vectors and a sum of w_i's of components for any of these vectors be $\sum_{i=1}^{n} w_i x_i = t$, where t is within the range $0 \leqq t \leqq M$. For the remaining w_i''s, $\sum_{i=1}^{n} w_i' x_i \geqq T - t$ must hold. Then $m_{T-t} \cdot n_t$ takes the value of 1 as both m_{T-t} and n_t are 1. Conversely, only when both $\sum_{i=1}^{n} w_i x_i \geqq t$ and $\sum_{i=1}^{n} w_i x_i \geqq T - t$ are satisfied, $m_{T-t} \cdot n_t$ takes the value of 1 and all input vectors validating these inequalities belong to S. Consequently, f_T and $\sum_{t=0}^{M} m_{T-t} \cdot n_t$ are simultaneously 1 for input vectors in S and simultaneously 0 for input vectors outside S. Thus expansion (33) is obtained.

An inductive synthesis of certain types of functions with n variables from a function of $(n-1)$ variables realized by a single element may be obtained as follows.

Corollary 3: If a function $f_t(x_1, \cdots, x_{n-1})$ of $(n-1)$ variables realized by an element with fixed coupling numbers is known for any possible threshold value t, a function of n variables

$$f_u(x_1, \cdots, x_{n-1}) + x_n \cdot f_v(x_1, \cdots, x_{n-1}) \tag{34}$$

can be realized with a single element, where between arbitrary thresholds u and v, a relation $u > v$ holds.

Proof: Set $w_i = 0$ $(i = 1, \cdots, n-1)$, $w_n \neq 0$ and $w_i' \neq 0$ $(i = 1, \cdots, n-1)$; $w_n' = 0$ in Theorem 7. Then we have $m_0 = 1$ and for t such that $M = w_n \geqq t \geqq 1$, $m_t = x_n$. Thus, rewriting n_T and n_{T-t} in (33) as f_T and f_{T-t}, we have $f_T(x_1, \cdots, x_{n-1}) + \sum_{t=1}^{M} x_n f_{T-t}$ $(x_1, \cdots, x_{n-1})$. As $f_{T-M} \geqslant f_{T-t}$ for $t < M$ holds from Theorem 5, this equation is simply

$$f_T + x_n \cdot f_{T-M}. \tag{35}$$

Though w_i''s $(i = 1, \cdots, n-1)$ are given, values of w_n and T may be optionally selected and therefore the stated function is realizable.

As a special case of the majority decision functions of (34), set $T = M$ and then $f(x_1, \cdots, x_{n-1}) + x_n$ is realizable. If we set $T = \sum_{i=1}^{n-1} w_i' + 1$, we have $x_n f(x_1, \cdots, x_{n-1})$ as $n_T = 0$ holds. Also for

$T = M + 1$ and $T = \sum_{i=1}^{n-1} w_i', f(x_1, \cdots, x_{n-1}) + (x_1 + x_2 \cdots + x_{n-1})x_n$ and $x_1 \cdots x_{n-1} + x_n f(x_1, \cdots, x_{n-1})$, respectively, are realizable.

4. FORMS OF FUNCTIONS REALIZABLE BY A SINGLE MAJORITY DECISION ELEMENT AND VARIOUS BOUNDS

It is an interesting problem to find how many functions of n variables can be realized by a single element, by adjusting coupling numbers and a threshold. Here upper and lower bounds are discussed since exact enumeration seems difficult.

Forms of Realizable Functions

Though enumeration of realizable types of functions by a single element gives its lower bound, simple types of functions will be discussed in this section. As shown already (5),

Theorem 8: A symmetric polynomial of degree d for n variables can be realized by a single majority decision element of structure $w_1 = w_2 = \cdots = w_n = 1$ and $T = d$.

The following function, where certain terms are eliminated from the above polynomial, is also realizable.

Theorem 9: A function in which a symmetric polynomial of the second degree for $m (< n)$ variables is eliminated from a symmetric polynomial of the second degree for n variables, can be realized by a single majority decision element. The number of such functions is $\sum_{m=3}^{n-1} \binom{n}{m} = 2^n - n - 2 - \binom{n}{2}$, where $n - 1 \geqq m \geqq 3$. (A function for $m = 2$ is a special case for the next theorem with $d = 2$.)

Theorem 10: A function where one term is eliminated from a symmetric polynomial of degree d for n variables can be realized by a single majority decision element. The number of such functions is $\binom{n}{d}$.

Proof: Theorems 9 and 10 will be proved. Eliminate a symmetric polynomial for m out of n variables from a symmetric polynomial of degree d for n variables. We get a realization as follows: Let α be the coupling numbers of the variables contained in the eliminated polynomial and $(\alpha + \beta)$ be the coupling numbers of the remaining $(n - m)$ variables, where α and β are positive integers, and let the threshold of the element be $T = \alpha m + 1$. Then a structure of the element is specified. First, since a sum of coupling numbers of m variables included in the eliminated polynomial is at most αm and does

not reach its threshold, a product of these m variables is not included in the function for the element. Next check a product of $(d-k)$ variables out of these m variables and k variables out of the remaining $(n-m)$ variables, where naturally $n > m \geqq d \geqq k \geqq 1$. A sum of coupling numbers of variables that are included in this product of degree d is

$$(\alpha+\beta)k + \alpha(d-k) = \alpha d + \beta k. \tag{36}$$

If

$$\beta = \alpha(m-d) + 1 \tag{37}$$

is chosen, Eq. 36 becomes $\alpha\{(m-d)k + d\} + k \geqq T = \alpha m + 1$ as $m - d \geqq 0$ and the sides are equal for $k = 1$, and then this value reaches its threshold. It remains to be shown that when any one variable is eliminated further from such a term, the sum of coupling numbers of the remaining variables does not reach the threshold. The largest sum of coupling numbers of variables of this term where one variable is excluded is

$$(\alpha+\beta)(d-1). \tag{38}$$

In order that this be smaller than T,

$$(d-1)\{\alpha(m-d+1)+1\} \leqq \alpha m \tag{39}$$

must hold. Then we have

$$\alpha\{m(d-2) - (d-1)^2\} + d - 1 \leqq 0. \tag{40}$$

This will be satisfied by setting

$$\alpha = d - 1 \tag{41}$$

for $m = d$ and particularly for $d = 2$ by setting $\alpha = 1$, irrelevant of m. Thus $\alpha + \beta$ and α are d and $d-1$, respectively, in the former case and m and 1 in the latter. The above proof is not valid for $d = 1$ but it is easy to see that Theorem 10 holds even for this case because the function is of the first degree.

Theorem 11: A function obtained from a symmetric polynomial of the second degree for n variables by eliminating a symmetric polynomial of the second degree for m variables and adding a symmetric polynomial of not less than the third degree for the same m variables, can be realized by a single majority decision element and this number is given by

$$\sum_{m=3}^{n-1} \binom{n}{m} (m-2), \tag{42}$$

where $n - 1 \geqq m \geqq 3$.

Theorem 12: A function in which a symmetric polynomial of the d^{th} degree for m variables contained in a symmetric polynomial of the d^{th} degree for n variables is replaced by a symmetric polynomial of the $(d+1)^{th}$ degree for the same m variables, can be realized by a single majority decision element. This number is given by

$$\sum_{m=d+1}^{n-1} \binom{n}{m}, \tag{43}$$

where $n - 1 \geqq m \geqq d + 1$.

Proof: Both theorems above will be proved together. α will be the coupling numbers of m variables and $\alpha + \beta$ the coupling numbers of $(n - m)$ variables and the threshold of the element will be $T = \alpha d + \alpha f + 1$, where α and β are positive integers and f a non-negative integer. With respect to the first m variables, any product of $(d + f + 1)$ variables of these has a sum of coupling numbers equal to $\alpha(d + f + 1)$ which is over the threshold, but any product of $(d + f)$ variables has a sum of coupling numbers which does not reach the threshold. Therefore a symmetric polynomial of $(d + f + 1)^{th}$ degree for m variables exists in the polynomial which the element represents. Next, a sum of coupling numbers corresponding to a product which consists of k of the remaining $(n - m)$ variables and $(d - k)$ of the first m variables, is $\alpha d + \beta k$, being calculated in a manner similar to Eq. 36; it always exceeds the threshold for $k \geqq 1$ if we choose

$$\beta = \alpha f + 1. \tag{44}$$

The largest sum of coupling numbers for a product in which a single variable is eliminated from this product is the value equal to (38). In order not to reach the threshold

$$(\alpha + \beta)(d - 1) \leqq \alpha d + \alpha f \tag{45}$$

must hold. Substitution of Eq. 44 gives

$$\alpha\{f(d - 2) - 1\} + d - 1 \leqq 0. \tag{46}$$

Then when $f = 0$ or $d = 2$ holds, (45) holds if we set $\alpha = d - 1$. Thus the theorems are proved. $\alpha + \beta$ and α are $f + 2$ and 1, respectively, in Theorem 11 and d and $d - 1$, respectively, in Theorem 12. It is easy to show similarly that the latter theorem holds even for $d = 1$.

An example will be shown. The symmetric polynomial of the second degree for four variables

$$x_1x_2 + x_1x_3 + x_2x_3 + x_1x_4 + x_2x_4 + x_3x_4 \tag{47}$$

can be realized by a single element. According to Theorem 9, make the following two functions by eliminating a symmetric polynomial of the second degree for $m = 2$, x_3x_4 and one for $m = 3$, $x_2x_3 + x_2x_4 + x_3x_4$, respectively, from this function:

$$x_1x_2 + x_1x_3 + x_2x_3 + x_1x_4 + x_3x_4 \tag{48}$$

and

$$x_1x_2 + x_1x_3 + x_1x_4. \tag{49}$$

Then both can be realized by a single element. Also, according to Theorem 10, any function made by eliminating any single term from (47) can be realized by a single element. According to Theorem 11 or 12, the function

$$x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3x_4 \tag{50}$$

in which a symmetric polynomial of the second degree for $m = 3$ in (47) is replaced by the third degree product, $x_2x_3x_4$, can be realized by a single element.

Theorem 13: A function in which a single variable in a symmetric polynomial of the d^{th} degree for n variables is replaced throughout by a product of any number of new variables can be realized by a single majority decision element.

Proof: The variable x_n is to be replaced by a product of m variables $y_1 \cdots y_m$. Multiply all the coupling numbers of $x_1, \cdots, x_{n-1}$ and T by m. Coupling numbers of $y_1, \cdots$, and y_m should be set to one. Realization of the function is then obvious.

For example, if we replace x_3 in $x_1x_2 + x_1x_3 + x_2x_3$ by x_3x_4, we have a realizable function

$$x_1x_2 + x_1x_3x_4 + x_2x_3x_4.$$

In an asymmetric polynomial, however, such replacement does not always give us realizable functions.

Lower Bound on a Number of the Majority Decision Functions

In this and the following sections, a number of the functions that can be realized by a single majority decision element will be discussed. First the symmetric polynomials in Theorem 8 will be considered. There are $\binom{n}{k}$ combinations of k out of n variables and the degrees of symmetric polynomials of k variables range from the first to the k^{th}. Negation of different sets of variables gives different functions. Thus the number of such functions is

$$\left| \sum_{k=1}^{n} 2^k \cdot k \cdot \binom{n}{k} \right| + 2, \tag{51}$$

including two functions that take constant values of 1 and 0. Summing the contents of the brackets, we have the following theorem.

Theorem 14: The number of symmetric polynomials for at most n variables that can be realized by a single majority decision element is

$$2n \cdot 3^{n-1} + 2, \tag{52}$$

including functions made by permutations and negations of variables.

This number is very small compared with 2^{2^n}, the number of functions of n variables. However, we shall show that the number of asymmetric polynomials is much greater, though it is still much smaller than 2^{2^n} for large n. Functions shown in Theorems 9 to 13 are asymmetric polynomials based on symmetric ones but their number is quite small. The structures of two types of asymmetric polynomials in the next theorem are entirely different from those above, but their number is large enough to deduce a reasonable lower bound on the number of functions. Forms of these functions in general are complex and it may be difficult to show them explicitly, but their differentiation is possible.

Theorem 15: A lower bound on the number of asymmetric polynomials for at most n variables that a single majority decision element can represent, is

$$\sum_{k=9}^{n} 2^k \cdot k! \cdot \left(2e - 1 - \frac{1}{(k-9)!}\right) + \\ + 2^n \binom{n}{[n/\log n]} ([n/\log n] - 1)^{n-[n/\log n]} \tag{53}$$

for $n \geqq 9$, where $[x]$ means the integer that is the nearest to but not smaller than x (this notation is different from the one used in mathematical text books) and where $\log n$ is the natural logarithm of n.

Proof: We shall discuss first the type of the functions characterized by having only a single product of T^{th} degree and other products of degree lower than T. Henceforth the threshold T of the element will be kept constant, by adjusting properly a coupling number and a value of the constant input. Now, keeping coupling numbers of T of n variables, for example, $x_1, x_2, \cdots, x_T$, as 1, independently vary the coupling numbers of the remaining $(n - T)$ variables through integral values from 2 to T. The function represented by this element always consists of a product of T^{th} degree $x_1 x_2 \cdots x_T$ and products of lower degree, each of which includes at least one of the other variables.

Since the T^{th} degree product term cannot be deleted from the function expression by including any other term as its factor, all functions for this T differ from those obtained by starting from different values of T. Let w_q be a coupling number of some variable x_q of the remaining $(n - T)$ variables, where $T \geqq w_q \geqq 2$. A product of this x_q and $(T - w_q)$ variables of $x_1, \cdots, x_T$ exists in this polynomial function and a sum of their coupling numbers takes a value of just T when all factor variables take the value of 1. For a different value of w_q, the number of variables in this product extracted from $x_1, \cdots, x_T$ differs. Now if we compare two functions for any two different sets of coupling numbers for variables x_{T+1} to x_n, we can find at least one different term such as the above. Since all possibilities produce different functions, the number of polynomials made in this way is

$$(T - 1)^{n-T}. \tag{54}$$

Selection of T variables and variation of T from 2 to $n - 1$ give polynomials of a number

$$\sum_{T=2}^{n-1} \binom{n}{T} (T - 1)^{n-T}. \tag{55}$$

Here the case $T = n$ is omitted since it represents a symmetric polynomial. Consequently the number of such polynomials for variables of a number equal to or less than n which can be represented by a single element is

$$\sum_{k=1}^{n} 2^k \binom{n}{k} \left[\sum_{T=2}^{k-1} \binom{k}{T} (T - 1)^{k-T} \right], \tag{56}$$

taking account of negation and permutation of the variables. Though evaluation of this formula is difficult, $(T - 1)^{k-T}$ has a larger effect on the magnitude of the formula than $\binom{k}{T}$ for large k and T. Examination of the largest effect of this leads to the conclusion that by choosing one term of the double sum:

$$T = [n/\log_e n], \qquad k = n, \tag{57}$$

formula 56 is bounded below by

$$2^n \binom{n}{[n/\log n]} ([n/\log n] - 1)^{n-[n/\log n]}, \tag{58}$$

where $[x]$ is an integer which is the nearest to but not less than x.

Next we shall discuss the second type of functions. These are obtained by changing the threshold value after assigning consecutive integers from 1 to n to coupling numbers of n variables as follows:

$w_1 = 1, \cdots, w_i = i, \cdots, w_n = n$. There are some cases where we can have explicit forms of these polynomials for some T's. First set $T = n$ and then a polynomial will be as follows:

$$
\begin{aligned}
& x_n + [x_{n-1}(x_{n-2} + \cdots + x_3 + x_2 + x_1) \\
& \quad + x_{n-2}(x_{n-3} + \cdots + x_3 + x_2) \\
& \quad + x_{n-3}(x_{n-4} + \cdots + x_3) \\
& \quad + \cdots \\
& \quad + \begin{cases} x_k x_{k-1} & \text{for odd } n,\ k = \dfrac{n+1}{2} \\ x_{k+1}(x_k + x_{k-1}) & \text{for even } n,\ k = \dfrac{n}{2} \end{cases} \\
& \quad + [\text{product terms of the degree equal to or more than the third, each of which includes at least one of the variables } x_1 \text{ to } x_{k-1} \text{ (for odd } n\text{, but to } x_k \text{ for even } n)].
\end{aligned} \tag{59}
$$

That is, there exists a product term of the first degree, x_n, a polynomial consisting of the second degree products that has some sort of regularity, and further for $n \geqq 6$, products of the degree equal to or more than the third. Now we shall determine when this polynomial consisting of the second degree products is invariant to permutation of the variables. In this polynomial the number of products, each of which includes x_j, and the number of products, each of which includes x_{n-j}, are j and $n - j - 1$, respectively, within a range of $1 \leqq j \leqq \dfrac{n-1}{2}$ for odd n. For even n these are j and $n - j - 1$, respectively, within a range of $1 \leqq j \leqq \dfrac{n}{2} - 1$ and furthermore there are $\left(\dfrac{n}{2} - 1\right)$ terms which include a single variable $x_{n/2}$. Therefore this polynomial cannot be the same if we permute the variables, some of which have different numbers of products, but there are possibilities of obtaining the same polynomials after permutations of those variables that have the same numbers of products counted above. Such variables are x_k and x_{k-1}, obtained by setting $j = \dfrac{n-1}{2}$ for odd n (the last term in the polynomial consisting of the second degree products in (59)) or x_{k-1} and $x_k (= x_{n/2})$, obtained by setting $j = \dfrac{n}{2} - 1$ for even n. Actually the following polynomials for $n = 4$, 5, and 6; $x_4 + x_3(x_2 + x_1)$, $x_5 + x_4(x_3 + x_2 + x_1) + x_3x_2$ and $x_6 + x_5(x_4 + x_3 + x_2 + x_1) + x_4(x_3 + x_2) + x_1x_2x_3$, are kept invariant for permutations of x_1 and x_2, x_2 and x_3, and x_2 and x_3, respectively. However for large m, if the third degree

terms in (59) are taken into account, this partial symmetry is broken, therefore giving different functions. That is, for odd n such that $n \geqq 7$, (that is, $k - 2 > 1$) $x_k x_{k-2} x_1$ exists in the third degree terms including x_k but the term $x_{k-1}x_{k-2}x_1$ is not there, so we can discriminate by the third degree terms two functions which are made by permutation of x_k and x_{k-1}. Also for even n such that $n \geqq 10$, (that is, $k - 2 > 2$) in the third degree terms including x_k there exist two terms $x_k x_{k-1} x_1$ and $x_k x_{k-2} x_2$ but the term $x_{k-1}x_{k-2}x_2$ is not there. Then permutation of x_k and x_{k-1} does not give the same function.

Consequently the number of functions given by permutation of n variables is

$$\begin{cases} \dfrac{n!}{2} & \text{for } n = 8 \text{ and } n \leqq 6 \\ n! & \text{for } n = 7 \text{ and } n \geqq 9. \end{cases} \tag{60}$$

Next, decreasing T lower than n gives functions that are different from the above and one another. The number of the first-degree terms in (59) increases, that is, $x_T + x_{T+1} + \cdots + x_n$ for T. With respect to variables x_1 to x_{T-1}, terms corresponding to ones of higher than the first degree in (59) appear in the functions. Then permutation of n variables gives many different functions, the number of which is obtained by

$$\left.\begin{aligned} &(T-1)!\binom{n}{n-(T-1)} = (T-1)!\binom{n}{T-1} \\ &\qquad\qquad \text{for } T = 7 \text{ and } T \geqq 9 \\ &\text{and} \\ &\tfrac{1}{2}\cdot(T-1)!\binom{n}{T-1} \qquad \text{for } T = 8 \text{ and } 3 \leqq T \leqq 6. \end{aligned}\right\} \tag{61}$$

Conversely increasing T starting at a value greater than n eliminates terms of the first degree. If we set $T = n + s$, a polynomial in the function, consisting of only the second-degree products, includes only variables from x_s to x_n for s such that $1 \leqq s \leqq n - 1$ and then the form of this polynomial is identical with that in (59), that is,

$$x_k(x_{n-1} + \cdots + x_s) + x_{n-1}(x_{n-2} + \cdots + x_{s+1}) + \cdots + \begin{cases} x_k x_{k-1} & \text{for odd } n+s,\ k = \dfrac{n+s+1}{2} \\ x_{k+1}(x_k + x_{k-1}), & \text{for even } n+s,\ k = \dfrac{n+s}{2}. \end{cases} \tag{62}$$

Also the function includes a generally complex polynomial consisting of terms of higher than the second degree. Though permutations of x_k

and x_{k-1} may give the same function, we can discriminate functions for $n \geqq 8$ by taking into account terms of higher than the second degree as in the case of $T = n$. Then the number of different functions given by permutation of n variables is not less than

$$(n - s + 1)! \binom{n}{s-1} = \frac{n!}{(s-1)} \tag{63}$$

for $n \geqq 8$, where $1 \leqq s \leqq n - 1$. Further increase of T gradually raises the lowest degree of products in the functions and makes their functional forms more complex.

Addition of three formulas (60, 61 and 63) over all possible T results in a number of the second type of functions given by permutation of n variables. If $n \geqq 9$, division by 2 is not required for $T \geqq 9$. A sum of (60) and (61), by use of $(n - g)! \binom{n}{g} = \frac{n!}{g!}$, results in

$$n! + \frac{n!}{2!} + \cdots + \frac{n!}{(n-8)!} \tag{64}$$

and a sum of (63) over s gives

$$n! + \left[n! + \frac{n!}{2!} + \cdots + \frac{n!}{(n-2)!} \right]. \tag{65}$$

The sum of these is greater than

$$n! \left[2e - \frac{1}{(n-9)!} - 1 \right]. \tag{66}$$

Calculations for cases of $n < 8$ (some of which are shown in Table II) are possible but somewhat complex. The number of functions given by taking into account negation of variables of a number equal to or less than n can be obtained by multiplying (66) by 2^n and adding them over integers not larger than $n (\geqq 8)$ but the number for n gives a good evaluation.

Last it may be shown as follows that the second type of functions differs from the first. In the second type of functions the polynomial consisting of only the second-degree products in each function has a partial symmetry, that is, it is invariant only to permutation of two specific variables (x_k and x_{k-1}). However, the first type of functions has one of the following cases: (1) a function does not have any second-degree product; (2) or, if it has, the polynomial consisting of the second-degree product in a function is invariant to permutation of more than three variables when $T \geq 3$. When $T = 2$, we have $x_1x_2 + (x_3 + \cdots + x_n)$ which does not appear in any of the second type functions.

However it should be noted here that even if repetition of some integers in a series of consecutive integers are repeated (like 433221), we still have functions with the same structure but with partial symmetries. If we take account of this synthesis, evaluation of the bound could be further improved. Table II includes functions of this structure.

Upper Bound on the Number of the Majority Decision Functions

First we shall state the following important result known in the theory of linear inequalities.

Lemma 1: Consider a set of consistent simultaneous linear inequalities (4)where $w_1, \cdots, w_n$ and T are regarded as variables. If the rank of its coefficient matrix is r, this set of inequalities is satisfied by a set of solutions obtained by equating r inequalities in (4) (10).

The following theorem is obtained with this Lemma.

Theorem 16: All functions of n variables that can be realized by a single majority decision element can be realized by properly selecting integers for coupling numbers of the input variables satisfying

$$w_i \leqq 2\left(\frac{n+1}{4}\right)^{(n+1)/2} \qquad (i = 1, 2, \cdots, n). \tag{67}$$

Proof: Assume that a given function is realized by a single element and then we have either the first type of an equation (4*a*) or the second type (4*b*), according to whether the function has the value 1 or 0 for each of 2^n sets of values of the variables. Then if this set of linear inequalities should be solved for $(n+1)$ variables $w_1, \cdots, w_n$ and T, Lemma 1 shows that if we choose and equate inequalities of a number $n+1$ out of the 2^n inequalities, their solutions satisfy all other inequalities. Writing these equalities as

$$\begin{aligned} &w_1 + \cdots + w_n - T = 0 \\ &\cdots \\ &w_1 + \cdots + w_n - T = -1 \\ &\cdots, \end{aligned} \tag{68}$$

we have

$$w_i = \frac{\Delta_i}{\Delta} \qquad (i = 1, 2, \cdots, n) \tag{69}$$

Here Δ is a coefficient determinant of (68) where all elements in the first to the n^{th} columns are either 1 or 0 and all elements in the $(n+1)^{\text{th}}$ column are -1. Δ_i is a determinant such that the i^{th} column of this determinant is replaced by the right sides of (68) consisting of 0's and -1's. Then divide all elements in the $(n+1)^{\text{th}}$ column by 2. Next add these $-\frac{1}{2}$ to other columns each of which consists of elements 1 and

0, and subtract these $-\frac{1}{2}$ from the column consisting of elements -1 and 0. Thus all elements in the determinant $2\Delta_i$ are $+\frac{1}{2}$ or $-\frac{1}{2}$. Application of the Hadamard's evaluation of a determinant leads to

$$|\Delta_i| \leq 2\left(\frac{n+1}{4}\right)^{(n+1)/2}. \tag{70}$$

As Δ is an integer not smaller than 1, then $w_i = \Delta_i/\Delta$ and T are rational numbers satisfying the inequalities. However if all w_i's and the threshold are multiplied by Δ, we have a logic element, with integral coupling numbers $\Delta \cdot w_i$, which still represents the same function, because Δ_i's are integers.

It may be interesting to compare the bound (67) with the largest coupling numbers required actually as follows.

n	*actual largest coupling number* (16, 17)	*bound* (67)
2	1	1
3	2	2
4	3	3
5	5	6
6	9	14
7	?	32

Theorem 17: An upper bound on the number of functions of variables of a number equal to or less than n that can be represented by a single majority decision element is

$$(2^n - 1)\left\{4\left[\frac{n+1}{4}\right]^{(n+1)/2} + 1\right\}^n. \tag{71}$$

Proof: As in the proof of Theorem 16 it was shown that the maximum of Δ_i is $2\left(\frac{n+1}{4}\right)^{(n+1)/2}$, a number of combinations of integral coupling numbers for n variables is the n^{th} power of this maximum. But as functions of less than n variables should be counted, $w_i = 0$ for all i's should be included and then the number is at most

$$\left\{2\left(\frac{n+1}{4}\right)^{(n+1)/2} + 1\right\}^n. \tag{72}$$

Since Theorem 6 showed that the number of possible values of T that give different functions is at most $2^n - 1$ excluding the two functions of 1 and 0 for each of the above combinations, the number of functions is at most the value given by (71), taking into account negation of variables. As this formula is counted with some redundancy, addition of two to this is not necessary in order to count the two functions of 1 and 0.

In order to explain the extremely large memory capacity of the human brain as compared with the number of neurons, 10^{10}, D. G. Willis (11) proposed a hypothesis that variation of coupling numbers of inputs gives a single neuron the capability to represent various complex functions, providing it a large memory capacity. Quantitative evaluation for this may be obtained from the above two theorems. As it is said that a single neuron may have inputs as large as several hundred, if we assume $n = 128$, a lower bound on the number of functions represented by a single neuron is 846 bits from (53), and its upper bound is about 4.1×10^4 bits from (71). For $n = 256$, a lower bound is about 1900 bits and an upper bound about 2×10^8 bits.

From Theorem 16, we have the following Corollary.

Corollary 4: All possible functions of n variables that can be realized by a single majority decision element can be realized with the integral values of total coupling number not larger than

$$4n\left(\frac{n+1}{4}\right)^{(n+1)/2} - 1. \tag{73}$$

Proof: If the threshold T is larger than $\sum_{i=1}^{n} w_i$, we have only the constant function that takes the value 0 permanently. Hence,

$$T \leqq \sum_{i=1}^{n} w_i \tag{74}$$

should hold. As the relation (3) holds for positive integral coupling numbers of variable inputs and the maximum of the total coupling number also occurs with a constant input of 0, the total coupling number is $2T - 1$ from (3). When w_i's are not integers, (74) and (67) hold even if $\Delta \cdot w_i$'s are regarded as new coupling numbers. We therefore now have (73).

5. CRITERION OF REALIZABILITY AND STRUCTURE DETERMINATION OF THE MAJORITY DECISION ELEMENT FOR A GIVEN FUNCTION BY LINEAR PROGRAMMING [4]

In order to realize a single majority decision element for a given function, we must determine the structure of the element so as to satisfy a set of the linear inequalities (4*a* and 4*b*) characterizing the function, but solutions cannot be uniquely obtained. For example if a

[4] After preparation of the manuscript of this paper, the authors found two related papers; R. McNaughton, "Unate Truth Functions," Appl. Math. and Statistics Lab., Tech. Rep. No. 4, Stanford University, October, 1957 (appeared in *Trans. IRE*, Vol. EC-10, No. 1, 1960); R. C. Minnick, "Synthesis of Linear-Input Logic by the Simplex Method," presented at the Sixth Annual Symposium on Computers and Data Processing of the Denver Research Institute, July, 1959 (*ibid.*). The authors believe that the present paper does not overlap with the first (except McNaughton's theorem 3.2 in the *Trans. IRE* edition) and may include details different from the second, such as the reduction processes of the number of inequalities.

EDITOR'S NOTE: This manuscript was submitted to the JOURNAL office July 27, 1960.

set of solutions is discovered, when it is multiplied by an arbitrary number larger than 1, this gives another set of solutions. Then we shall add a restricting condition to minimize $K = w_1 + w_2 + \cdots + w_n + w_e$, the total coupling number of the element shown in (1). This restriction is actually important since it has economic and engineering merits. That is, the minimization of coupling amplitude among elements may result in a reduction of back couplings among elements and an increase of the number of outputs that can be obtained from a single element. These factors imply better stability of the circuits (15). When couplings among elements are realized by wiring as in the case of parametrons, this condition corresponds to minimizing the total sum of wiring turns. When elements are coupled with resistors as in a certain type of circuitry with the Esaki diodes, this corresponds to a maximization of resistor values. If it is desirable for the resistor values to be within a certain range, we must multiply a set of solutions by a number of proper magnitude or add inequalities representing a range for the resistor values, to the inequalities (4*a* and 4*b*).

Adding the minimization condition to the set of the inequalities (4*a* and 4*b*) in this way, determination of the structure of a majority decision element for a given function can be reduced to a typical problem in linear programming that gives us at the same time a criterion of its realizability, as shown later.

Since there are duality and inclusion relations between two majority decision functions, one with a constant input 0 and the other with a constant input 1, but with the same coupling numbers of variable and constant inputs as shown in the section on *Duality of the Majority Decision Functions* (p. 387), determination of the structure of the element can be simplified as follows: First check whether the given function can be realized by an element without a constant input, by finding whether the function is self-dual. Next, if this is not the case, compare the function with its dual and choose the smaller one (see p. 387). If it is realized by a single element, the element must have a constant input 0 and the other larger function can be realized by an element of the same structure except with a constant input 1. Consequently we shall henceforth only consider realizations of elements with no constant input or with a constant input 0. If there is no inclusion relationship between the given function and its dual, this function is not realizable, as seen from Corollary 2. Of course other various necessary conditions already discussed can be used to reject unrealizable functions.

Now for simplicity, let us assume that all coupling numbers are integers and that K is odd. A more general case where some of them are non-integers will be discussed at the end of this section.

As when all coupling numbers are integers the relation (3) between T and K can be used, the problem is simplified, excluding w_e.

Problem 1: When all coupling numbers w_i, and a threshold T to be found, are assumed as positive integers, determination of the structure of a majority decision element for a given function is reduced to a linear programming problem; the objective function, T, is to be minimized under constraints of a set of the inequalities (4*a* and 4*b*). (The objective function is actually $2T - 1$ but constants -1 and 2 can be ignored.)

By the duality theorem in the theory of linear programming this problem can be converted into the following Problem 2 where the objective function is to be maximized. Solutions for this problem will be given by use of the simplex method and will simultaneously give solutions to the Problem 1. In this case there is no guarantee that the $w_1, w_2, \cdots, w_n$ and T obtained are integers, but these will always be rational numbers since all coefficients in the set of inequalities are integers. When these happen to be non-integers, integral solutions can be obtained by the method developed recently by R. Gomory (12), but we shall not discuss this in this paper.

Problem 2: Problem 1 is dual to the problem: under the constraints

$$\sum_{j=1}^{r} v_j x_{ij} - \sum_{j=r+1}^{m} v_j x_{ij} \leqq 0 \qquad (i = 1, 2, \cdots, n)$$

$$-\sum_{j=1}^{r} v_j + \sum_{j=r+1}^{m} v_j \leqq 1 \tag{75}$$

and

$$v_j \geqq 0 \qquad (j = 1, 2, \cdots, m), \tag{76}$$

the objective function

$$\sum_{j=r+1}^{m} v_j \tag{77}$$

is to be maximized.

The realization problem is thus converted into the regular type of linear programming problem, which can be solved by making tableaux of the simplex method. However since coefficients in condition (75) are only ± 1 and 0, it is generally a case of degeneracy and then in making tableaux the perturbation must be taken into account. By this perturbed version of the simplex method, if a given function is to be realized by a single element, w_i ($i = 1, 2, \cdots, n$) and T can be determined as optimum solutions and furthermore if it is not, that is, if Problem 1 does not have solutions, the non-existence of solutions is shown in the last tableau in Problem 2.

The number of inequalities included in (4*a* and 4*b*) is just 2^n, since there are n variables. Even for small values of n this number becomes

too large to be handled by electronic digital computers and solution by the simplex method may appear impossible. However, by the following processes we can cut it down greatly.

Process A: Choose any one of the inequalities of (4*a*). If it is satisfied by variables $w_1, \cdots, w_n$ and T, any other inequality in (4*a*) such that any other terms are added to the chosen inequality, if such an inequality exists, is also satisfied. Then the latter inequality can be eliminated as a redundant one from (4*a*). Elimination of redundant inequalities in this way leaves inequalities each of which has the least number of terms. These will be called "the set of the minimal-term inequalities."

With respect to the inequalities of (4*b*), if there exists an inequality each of whose terms appears in some other inequality that may itself have more terms than the original, then the original inequality can be eliminated as redundant from (4*b*). This elimination process of redundant inequalities leaves inequalities each of which has the largest number of terms. These will be called "the set of the maximal-term inequalities."

Now let us consider the Boolean lattice constructed on the partial ordering defined by

$$X \supseteq X', \tag{78}$$

for two sets, X and X', each of which consists of numbers from 1 to n. As easily seen, the maximal-term inequalities correspond to the generators of the principal ideals in this Boolean lattice and the minimal-term inequalities correspond to the generators of the principal dual ideals. Each element of the lattice must belong to one of these ideals or dual ideals.

The process of extracting the minimal-term inequalities from those of (4*a*) is equivalent simply to an application of the deletion law described in the paragraph after Definition 2, that any term including some other term as its factor in the function expression can be deleted. Then we can write down the minimal-term inequalities directly from the given function as each term in the polynomial expression for the function, that is, each prime implicant has one-to-one correspondence with each of the minimal-term inequalities. Now there is a question of whether we can write down the maximal-term inequalities from this polynomial. The procedure is as follows.

Process B: Make the polynomial dual to the given polynomial and then eliminate all redundant terms by application of the deletion law. Next, for each term included in this dual polynomial make a product consisting of only variables that are not included in the term, and now all such products specify the maximal-term inequalities, each product corresponding to each inequality.

This is easily seen from the fact that negation of the original function specifies the maximal-term inequalities and it is equivalently expressed by its dual function with negated inputs.

For example, when the function $x_1x_2 + x_1x_3 + x_2x_3 + x_1x_4$ is given, the minimal-term inequalities corresponding to each term will be as follows,

$$\begin{aligned} w_1 + w_2 \qquad\quad - T &\geqq 0 \\ w_1 \qquad + w_3 \qquad - T &\geqq 0 \\ w_2 + w_3 \qquad - T &\geqq 0 \\ w_1 \qquad\quad + w_4 - T &\geqq 0. \end{aligned} \tag{79}$$

Figure 1 shows the Boolean lattice of the minterms where the lattice elements are indicated by suffixes of the variables. The inequalities in

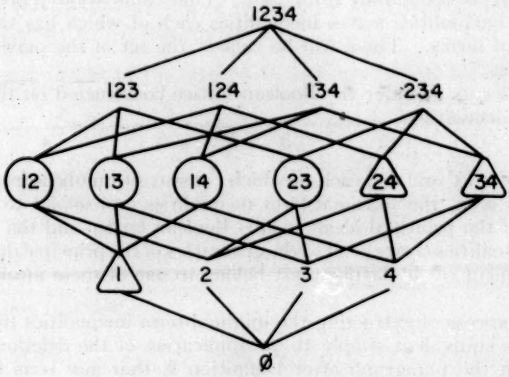

FIG. 1. Lattice for four variables.

(79) correspond to the lattice elements with circles. Next, as its dual function is $x_1x_2 + x_1x_3 + x_2x_3x_4$, according to the process above we have a sum of three terms x_3x_4, x_2x_4 and x_1 corresponding to the respective terms of the function. The lattice elements with the triangles in Fig. 1 indicate these terms and, together with the lattice elements with the circles, bisect the whole lattice, where ϕ indicates the lower bound of the lattice. The maximal-term inequalities are obtained from these terms as follows:

$$\begin{aligned} \qquad\quad - w_3 - w_4 + T &\geqq 1 \\ - w_2 \qquad - w_4 + T &\geqq 1 \\ - w_1 \qquad\quad + T &\geqq 1. \end{aligned} \tag{80}$$

These processes reduce $2^4 = 16$ inequalities to only seven inequalities in (79) and (80).

The number of these inequalities may be, in general, reduced further by the following process of comparing coupling numbers of variables indirectly, though of course we shall get exactly identical solutions without this process, if solutions exist.

Process C: Choose any two terms A and B in the given original function and compare them in the following way. Let $f_{A/B}$ be the function obtained by assigning values of 1 to the variables included only in the term A of the original function and values of 0 to the variables included only in the other term B. This is a function of the residual variables. Next, let $f_{B/A}$ be the function obtained by assigning 0 to the former variables and 1 to the latter. If $f_{A/B} \supset f_{B/A}$, eliminate an inequality corresponding to term A from the minimal-term inequalities and, if $f_{A/B} \subset f_{B/A}$, eliminate an inequality corresponding to B. If neither holds, do not eliminate either. Repeat this process over the remaining terms. Finally, when any inequality cannot be further eliminated for selection of any two terms, make inequalities corresponding to these terms remaining. In many cases, this process eliminates further inequalities from the minimal-term inequalities.

The reason why elimination of inequalities is possible by this process is explained by the following considerations. A set of feasible solutions of Problem 1 constitutes a bounded polyhedron and the optimum solutions correspond to one of its vertices. Then the optimum solutions are obtained by solving a set of simultaneous linear equalities obtained by removing the greater-than condition from the inequalities chosen properly out of the set of the inequalities (4*a* and 4*b*) (**13**, **14**). However, $f_{A/B} \supset f_{B/A}$ means that a sum of coupling numbers of variables included only in A is larger than that included only in B. Then as it means that the left side of the inequality corresponding to A is larger than that corresponding to B, if only an equality "$=$" is to hold in the latter inequality "$\geqq$," only an inequality "$>$" but not an equality "$=$" should hold in the former inequality. Consequently the former inequality corresponding to A is unnecessary, leaving only the equated constraint corresponding to B in the set of the simultaneous linear equations. However it is to be noted here that since we cannot in general eliminate all of the redundant inequalities, we do not know which of the remaining inequalities should be equalities and consequently we must solve a set of inequalities, as $\geqq$ and not $=$. But when the function is a function of truly n variables, that is, when there are just $n + 1$ constraints, all of them can be equated.

We shall explain the above process on the function used earlier: $x_1x_2 + x_1x_3 + x_2x_3 + x_1x_4$. First compare x_1x_2 and x_2x_3. A variable included only in the former is x_1 and one included only in the latter is x_3. A function obtained by setting $x_1 = 1$ and $x_3 = 0$ is $x_2 + x_4$ and

one obtained by setting $x_1 = 0$ and $x_3 = 1$ is x_2. As $x_2 + x_4 \supset x_2$, an inequality corresponding to x_1x_2 can be eliminated. Next compare x_1x_3 and x_2x_3. A function given by setting $x_1 = 1$ and $x_2 = 0$ is $x_3 + x_4$ and one given by $x_1 = 0$ and $x_2 = 1$ is x_3. As $x_3 + x_4 \supset x_3$ holds, an inequality corresponding to x_1x_3 can be eliminated. As to the residual two terms, x_2x_3 and x_1x_3, such comparison of functions is not possible and then only inequalities corresponding to these two remain.

Process D: Process *C* should be again applied, comparing products obtained in Process *B*, except that inequalities corresponding to term *B* instead of *A* should be eliminated from the maximal-term inequalities, if $f_{A/B} \supset f_{B/A}$. The reason for this is the same as in the previous case, except that coefficients of w_i's are negative.

For the above example, Process *D* should be applied on the original function $x_1x_2 + x_1x_3 + x_2x_3 + x_1x_4$. By Process *B*, we had x_3x_4, x_2x_4 and x_1. For example, compare x_2x_4 and x_3x_4. A function given by setting $x_2 = 1$ and $x_3 = 0$, in the original function, is x_1 and a function given by setting $x_2 = 0$ and $x_3 = 1$ is x_1, identical to the first function. Also, for any two terms, we have the identical function. Then no inequality should be eliminated.

Process E: The above processes covered Problem 1. We must consider the dual Problem 2 with the eliminated set of inequalities after these processes. Then the number of inequalities in (75) is $n + 1$ but the number of variables v_j's is in general greatly reduced from 2^n.

A Numerical Example: Applications of the above Processes *A*–*E* on the inequalities (4*a* and 4*b*) for the function $x_1x_2 + x_1x_3 + x_2x_3 + x_1x_4$ reduce 16 inequalities to the following five:

$$\begin{aligned} w_2 + w_3 - T &\geqq 0 \\ w_1 + w_4 - T &\geqq 0 \\ -w_3 - w_4 + T &\geqq 1 \\ -w_2 - w_4 + T &\geqq 1 \\ -w_1 + T &\geqq 1. \end{aligned} \tag{81}$$

This is converted by Process *E* into the problem of finding the optimum values for $v_1, \cdots, v_5$ such that under constraints

$$\begin{aligned} v_2 - v_5 &\leqq 0 \\ v_1 - v_4 &\leqq 0 \\ v_1 - v_3 &\leqq 0 \\ v_2 - v_3 - v_4 &\leqq 0 \\ -v_1 - v_2 + v_3 + v_4 + v_5 &\leqq 1 \end{aligned} \tag{82}$$

TABLE I.—*Simplex Tableaux.*

		$j \rightarrow$	0	1	2	3	4	5	6	7	8	9	10
		$u_i \rightarrow$	0	0	0	0	0	0	0	0	1	1	1
$i \downarrow$	u_i	variable	c_i	λ_1	λ_2	λ_3	λ_4	λ_5	v_1	v_2	v_3	v_4	v_5
1	0	λ_1	0	1					0	1	0	0	−1
2	0	λ_2	0		1				1	0	0	−1	0
3	0	λ_3	0			1			1	0	−1	0	0
4	0	λ_4	0				1		0	1	−1	−1	0
5	0	λ_5	1					1	−1	−1	1	1	1
		$z_j - u_j$	0	0	0	0	0	0	0	0	−1	−1	−1
1	0	λ_1	0	1					0	1	0	0	−1
2	0	λ_2	0		1				1	0	0	−1	0
3	0	λ_3	1			1		1	0	−1	0	1	1
4	0	λ_4	1				1	1	−1	0	0	0	1
8	1	v_3	1					1	−1	−1	1	1	1
		$z_j - u_j$	1	0	0	0	0	1	−1	−1	0	0	0
1	0	λ_1	0	1					0	1	0	0	−1
6	0	v_1	0		1				1	0	0	−1	0
3	0	λ_3	1			1		1	0	−1	0	1	1
4	0	λ_4	1		1		1	1	0	0	0	−1	1
8	1	v_3	1		1			1	0	−1	1	0	1
		$z_j - u_j$	1	0	1	0	0	1	0	−1	0	−1	0
7	0	v_2	0	1					0	1	0	0	−1
6	0	v_1	0		1				1	0	0	−1	0
3	0	λ_3	1	1		1		1	0	0	0	1	0
4	0	λ_4	1		1		1	1	0	0	0	−1	1
8	1	v_3	1	1	1			1	0	0	1	1	0
		$z_j - u_j$	1	1	1	0	0	1	0	0	0	−1	−1
7	0	v_2	0	1					0	1	0	0	−1
6	0	v_1	1	1	1	1		1	1	0	0	0	0
9	1	v_4	1	1		1		1	0	0	0	1	0
4	0	λ_4	2	1	1	1	1	2	0	0	0	0	1
8	1	v_3	1	1	1			1	0	0	1	0	0
		$z_j - u_j$	2	2	1	1	0	2	0	0	0	0	−1
7	0	v_2	2	2	1	1	1	2	0	1	0	0	0
6	0	v_1	1	1	1	1		1	1	0	0	0	0
9	1	v_4	1	1		1		1	0	0	0	1	0
10	1	v_5	2	1	1	1	1	2	0	0	0	0	1
8	1	v_3	1	1	1			1	0	0	1	0	0
		$z_j - u_j$	4	3	2	2	1	4	0	0	0	0	0
			↑ T	↑ w_1	↑ w_2	↑ w_3	↑ w_4	↑ T					

and

$$v_j \geqq 0 \qquad (j = 1, 2, \cdots, 5), \tag{83}$$

the objective function

$$v_3 + v_4 + v_5 \tag{84}$$

should be maximized.

Now proceed with the solution of this problem using the simplex tableaux, taking into account the perturbation, as shown in Table I. Finally, if the problem has solutions, the structure of the majority decision element to realize the given function will be obtained. In this table the row "u_j" shows coefficients of the terms in the objective function (84). And $z_j = \sum_i u_i b_{ij}$ where b_{ij} is an entry in the i^{th} column and the j^{th} row. $\lambda_1, \lambda_2, \cdots$ and λ_5 are slack variables and the column "c_i" shows the right-side constants of the constraints.

As shown in the last row of the optimum tableaux, the optimum solutions that minimize the total coupling number as $2T - 1 = 7$, are $w_1 = 3$, $w_2 = w_3 = 2$, $w_4 = 1$ and $T = 4$, and this structure certainly represents a given function.

Process F: The tableaux are made as follows. First choose any entry, $z_k - u_k$, with a negative sign and with the largest absolute value in the row labeled as "$z_j - u_j$," and then find $b_{\alpha k}$'s with a positive sign in the column to which the above $z_k - u_k$ belongs. Since generally there are several $b_{\alpha k}$'s, calculate $\frac{1}{b_{\alpha k}}(c_\alpha + \sum_j b_{\alpha j}\epsilon^j)$ for each of such $b_{\alpha k}$'s, where c_α is the entry in the column labeled as "c_i" and where the summation $\sum_j$ is over the entries in the columns of the slack variables ($j = 1, 2, \cdots, 5$ in Table I) and where ϵ is a sufficiently small positive number. We can find only one $b_{\alpha^* k}$ such that the above formula has the minimum value. Next, replace all entries $b_{\alpha^* j}$ in the $\alpha^{*\text{th}}$ row by $b'_{\alpha^* j} = b_{\alpha^* j}/b_{\alpha^* k}$, except that the entry in the column labelled "variable," and in the $\alpha^{*\text{th}}$ row, is replaced by the symbol in the k^{th} column and in the row labelled "variable" and that the entry in the column labelled "u_i" and in the $\alpha^{*\text{th}}$ row is replaced by a number in the row labelled "u_j" and in the k^{th} column. Then, replace all the other entries in the k^{th} column by 0 and the other b_{ij}'s in the tableau by $b_{ij} - b'_{\alpha^* j} \cdot b_{ik}$. Finally, calculate $z_j - u_j$. Thus, the next tableau was obtained.

For the next tableaux repeat the above process which is one usually used in the simplex method.

In the optimum tableau we shall finally have the following three cases. Here it is important to note the following: If we apply the simplex method on a set of inequalities reduced by Processes C and D, there is no guarantee for solutions obtained in the following cases (i) and (ii) to be solutions for the original Problem 1, too. Therefore we

must check whether the solutions obtained satisfy a set of inequalities obtained by applying only Processes A and B, or alternatively, whether the coupling weights and threshold obtained construct a given function. If there is no solution as in the following case (iii), the original Problem 1 also has no solution. However, when the simplex method is used after applying only Processes A and B, no such checking procedure is required.

(i) Consider the columns of variables (the 6th to 10th columns in Table I). If $z_j - u_j > 0$ holds for all columns except ones (adopted as the bases) where all entries except a single one are replaced by 0 in the above process, we have a single set of the optimum solutions for Problem 1 given by $(z_j - u_j)$'s in the columns of the slack variables.

(ii) If $z_j - u_j = 0$ holds for some columns (in columns other than the bases), we have other sets of the optimum solutions than the one in the optimum tableau.

(iii) If we cannot find any b_{ak} with a positive value in Process E, in other words, if only zeros or negative values are found in the column, the solutions for Problem 2 are called "infinite" because the objective function can be made arbitrarily large. In the original Problem 1 this means that there exist no solutions. When a given function cannot be represented by a single majority decision element, the last tableau always takes this form.

In the last tableau, if some slack variables remain in the bases, coupling numbers corresponding to them become 0. This is a case where a given function can be represented with fewer variables than those apparently required. It is interesting to note that if the coupling numbers and the threshold are not 0, the inequalities corresponding to these in (75) should hold with equalities "=" only (14).

So far no guarantee has been obtained that solutions $w_1, \cdots, w_n$ and T obtained by the above simplex method will always be integers. However, as mentioned later, we had observed that the optimum solutions w_i's and T are integers for any functions with variables of a small number n, though non-integers appeared in intermediate tableaux in some cases.

In the general case where there is a possibility that such non-integers may be included among coupling numbers of the inputs, it may not be reasonable to solve Problem 1 directly, where the objective function T was set up under the assumption that all coupling numbers are integers. Then we must solve the problem where the objective function $K = w_1 + \cdots + w_n + w_c$ should be minimized under the constraints (4a and 4b). However, in this case, T is included only in the constraints while w_c is included only in the objective function. Then a relation among T and $w_1, w_2, \cdots, w_n$ and w_c, should be added to the constraints. To avoid this, let us start from the following expression. An equivalent of (2a and 2b) in Definition 2 of the majority decision

elements, where w_e is used instead of T is: the output value of the majority decision element is

$$1, \text{ if } \sum_{i=1}^{n} w_i(2x_{ij} - 1) + (2x_e - 1)w_e > 0$$
$$0, \text{ if } \sum_{i=1}^{n} w_i(2x_{ij} - 1) + (2x_e - 1)w_e < 0, \tag{85}$$

where $x_{1j}, \cdots, x_{nj}$ are components of the j^{th} input vector ($j = 1, \cdots, 2^n$) and x_e is the value of the constant input.

Problem 3: Determination of the structure of a majority decision element for a given function is reduced to the linear programming problem: the objective function $K = w_1 + \cdots + w_n + w_e$ is to be minimized under constraints of a set of inequalities (85) where equalities "=" are included in all inequalities after knowing the minimum of the left sides of the first inequalities and the maximum of those of the second.

Here it is important to note that if the inequalities in (85) do not include equalities "=," then arbitrarily small w_i's and w_e can be solutions. This is an essential point in this minimization problem. Then the minimum ϵ_1 of the left sides of the first inequalities and the maximum ϵ_2 of those of the second inequalities should be known or predetermined. These give a margin in the performance of the element. Choose the smaller one out of ϵ_1 and $|\epsilon_2|$ and write it as ϵ, that is, $\epsilon = \text{Min } (\epsilon_1, |\epsilon_2|)$. All inequalities in (85) can be made to include equalities "=," by normalizing the inequalities by dividing these by ϵ. Therefore, Problem 3 may be used instead of Problem 1 and the simplex method can be applied again. Exactly the same processes as those used on Problem 1 can be applied here, as easily seen. This form of the problem may be more basic since we are working directly on w_e without T, but may be somewhat complex.

However, it is important in this case to see that if we introduce a parameter T characterizing the majority decision element given by $T = \frac{1}{2}(K + 1)$ after this normalization, the constraints in Problem 3 coincide with those in (4*a* and 4*b*) and the objective function also will be T. Consequently, Problem 3 coincides formally with Problem 1 and then even in the general case where non-integral coupling numbers may be included, it may be sufficient for us to solve Problem 1 formally.

Here it should be noted that in Problem 1 the relation between T and w_e, (3) is not included. The minimization of the threshold T only, as the objective function, without taking into account relation (3), implies intuitively minimization of sums of coupling numbers which do not reach the threshold.

TABLE II.—*Numbers of Functions.*

n	2^{2^n}	Number of Positive Functions of up to n Variables	Number of Majority Decision Functions of n Variables without Negation		Number of Majority Decision Functions of up to n Variables without Negation	Number $N(n)$ of Majority Decision Functions of up to n Variables with Permutation and Negation	Theoretical Evaluation of $N(n)$	
			Representative Functions*	Functions Counted with Permutation			Lower Bound	Upper Bound
1	4	3	1	1	3	4	4	/
2	16	6	2	2	6	14	14	27
3	256	20	5	9	20	104	104	875
4	65536	168	17	96	150	1882	1882	3.6×10^4
5	4294967296	7581	92	2690	3287	94572	32172	1.1×10^7
6	18 446744073 709551616	7828354	994	226360	244158	15028134	/	3.7×10^{10}

* Functions identical by permutations and negations of variables are counted into a single representative function.

6. MAJORITY DECISION FUNCTIONS OF UP TO SIX VARIABLES

In order to check various properties of the majority decision functions described so far, the Parametron Computer MUSASINO-1 (15) was used to find all majority decision functions of up to six variables and the structures of elements to realize these, based on Problems 1 and 2.

Majority Decision Functions of a Few Variables and Their Number

All majority decision functions of up to six variables are classified by permutations of variables, and negations of the variables. A representative function in each class such that $w_1 \geqq w_2 \geqq w_3 \geqq w_4 \geqq w_5 \geqq w_6 \geqq 0$ are obtained (16, 17). The number of these functions is too great to be shown here.

These functions were checked both by the simplex method and the combination method, where the functions are discovered by giving all possible integer values to the inputs.

Table II lists the numbers of functions enumerated from several different viewpoints (18).

(i) The number of general Boolean functions of up to n variables, 2^{2^n}, is shown for reference in the second column of Table II.

(ii) The number of the positive functions of up to n variables, those that can be realized without negation of the variables. These may generally not be realizable by a single majority decision element. The numbers were quoted from Birkhoff's book (19). The majority decision functions constitute only a subset of these.

(iii) The number of the majority decision functions of exactly n variables without negations of the variables. This is divided into two cases. The first case shows the number of the majority decision functions, representatives of equivalent classes, each of which consists of functions identical by permutations and negations of the variables. For example, $x_1x_2 + x_3$, $x_1x_3 + x_2$ and $\bar{x}_1 \cdot \bar{x}_2 + x_3$ belong to the same class represented by $x_1 + x_2x_3$. The second case shows the number of the functions in the first case, taking into account only permutations of variables. The functions in the second case are not enumerable unless the forms of the functions representing the above classes are known, because there are partially or totally symmetric functions. For example, for $n = 2$, we have two, counting $x_1 + x_2$ and x_1x_2.

(iv) The number of the majority decision functions of up to n variables, without negations of variables. For example, for $n = 2$ we have six, counting 0, 1, x_1, x_2, $x_1 + x_2$ and x_1x_2.

(v) The number $N(n)$ of the majority decision functions of up to n variables, taking into account permutations and negations of variables. If this is compared with 2^{2^n}, we can see how small a part of 2^{2^n} this occupies.

(vi) Theoretical upper and lower bounds on $N(n)$ shown in Section 4. The lower bound shows the number of functions constructed similarly to

the methods in Theorems, 8, 13, 14, and 15. (Cf. the paragraph just preceding *Upper Bound on the Number of Majority Decision Functions*, p. 400.) As the bounds in Section 4 were aimed for cases of large n, these approximations are not good here.

The ratio of the number of majority functions to that of the general functions or to that of the positive functions approaches zero as n increases.

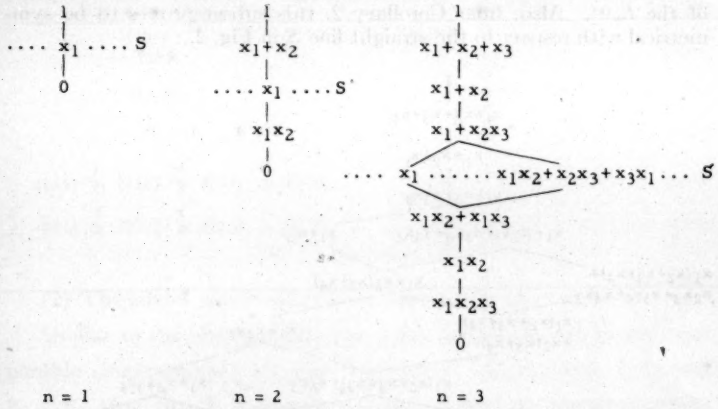

FIG. 2. $L_n^{(1)}$. $(w_1 \geq w_2 \geq w_3)$

Lattices of the Majority Decision Functions

Though there may be various ways to classify the majority decision functions, a classification based on the ordering among the input variables and that among the input vectors in Section 3 will be discussed here.

(1) The lattice based on ordering among the variables.

If an ordering $x_i \gtrsim x_j$ between any two variables x_i and x_j defined for a function $f_1(x_1, \cdots, x_n)$ holds also for another function $f_2(x_1, \cdots, x_n)$, we can define an ordering between x_i and x_j for the functions $f_1 + f_2$ and $f_1 \cdot f_2$ where $x_i \gtrsim x_j$ still holds. Therefore, the whole set of functions that defines ordering among the input variables, for example, $x_1 \gtrsim x_2 \gtrsim \cdots \gtrsim x_n$, clearly constitutes a distributive lattice with respect to the operations "+" and "·." This lattice will be denoted as $L_n^{(1)}$.

Though functions constituting the above lattice are not always the majority decision functions, it was ascertained that such functions are

confined to the majority functions only when the number of variables does not exceed five. Then the whole set of majority decision functions such that $w_1 \geqq w_2 \geqq \cdots \geqq w_n$, coincides with $L_n^{(1)}$ for $n \leqq 5$.

The Hasse diagrams of the $L_n^{(1)}$'s for $n \leqq 4$ are shown in Figs. 2 and 3. As shown in Theorem 5, when we increase the threshold of an element with fixed coupling numbers of variable inputs, the functions get smaller, as $f_0 \supseteq f_1 \supseteq \cdots \supseteq f_N$, where f_i is a function with the threshold i. A series of such functions $f_0, f_1, \cdots, f_N$ is a subset in a chain of the $L_n^{(1)}$. Also, from Corollary 2, this subset proves to be symmetrical with respect to the straight line S in Fig. 2.

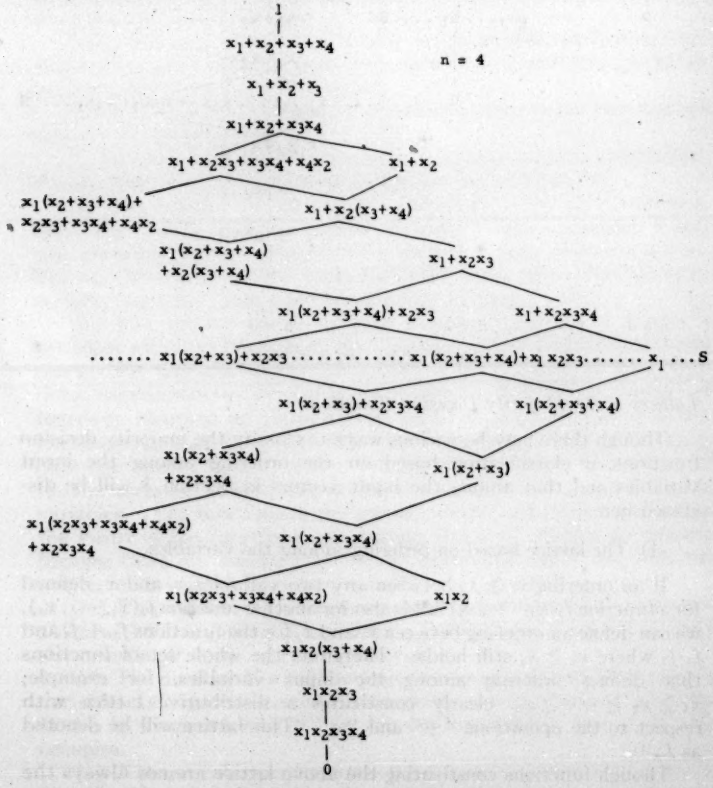

FIG. 3. $L_4^{(1)}$. $(w_1 \geq w_2 \geq w_3 \geq w_4)$

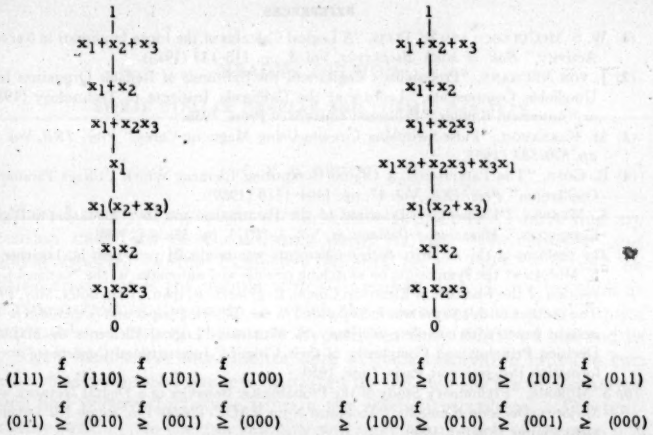

FIG. 4. Examples of $L_3^{(2)}$.

(2) The lattice based on orderings among the input vectors.

Similar to the above, if any two input vectors Ξ_i and Ξ_j are comparable simultaneously for two functions f_1 and f_2, as $\Xi_i \overset{f_1}{\geq} \Xi_j$ and $\Xi_i \overset{f_2}{\geq} \Xi_j$, then also $\Xi_i \overset{f_1+f_2}{\geq} \Xi_j$ and $\Xi_i \overset{f_1 \cdot f_2}{\geq} \Xi_j$. Therefore, the whole set of the functions which arrange all the input vectors $\{\Xi_i\}$ in a certain order proves to constitute a distributive lattice with respect to the operations "+" and "·."

We can construct various lattices according to what ordinal relations are given among $\{\Xi_i\}$. However, we cannot establish the relations on $\{\Xi_i\}$ in a completely arbitrary way so that these can be defined by some functions.

Now let $L_n^{(2)}$ be a special lattice such that orderings $\Xi_1 \overset{f}{\geq} \Xi_2 \overset{f}{\geq} \cdots \overset{f}{\geq} \Xi_{2^n}$ are established and $x_1 \gtrsim x_2 \gtrsim \cdots \gtrsim x_n$ hold. $L_n^{(2)}$ is a sublattice of $L_n^{(1)}$. It was ascertained in a range of $n \leqq 5$ that all functions included in $L_n^{(2)}$ are majority decision functions and some chains of $L_n^{(1)}$.

Figure 4 shows examples of $L_3^{(2)}$. These exhaust all chains of $L_3^{(1)}$ but in the cases of $n = 4$ and 5, not all chains of $L_n^{(1)}$ are included in $L_n^{(2)}$.

Acknowledgment

The authors are indebted to C. C. Elgot and J. B. Wiener for their help in revising the manuscript.

REFERENCES

(1) W. S. MCCULLOCH AND W. PITTS, "A Logical Calculus of the Ideas Imminent in Nervous Activity," *Bull. of Math. Biophysics*, Vol. 5, pp. 115–133 (1943).

(2) J. VON NEUMANN, "Probabilistic Logics and the Synthesis of Reliable Organisms from Unreliable Components," Lecture at the California Institute of Technology (1952), or "Automata Studies," Princeton University Press, 1956.

(3) M. KARNAUGH, "Pulse-Switching Circuits Using Magnetic Cores," *Proc. IRE*, Vol. 43, pp. 570–583 (1955).

(4) E. GOTO, "The Parametron, a Digital Computing Element Which Utilizes Parametric Oscillation," *Proc. IRE*, Vol. 47, pp. 1304–1316 (1959).
S. MUROGA, "Elementary Principle of the Parametron and Its Application to Digital Computers," *Magazine of Datamation*, Vol. 4, No. 5, pp. 31–34 (1958).

(5) The problem of the majority decision elements was originally presented in Japanese by S. Muroga at the symposium on switching circuits and automata, at the National Convention of the Institute of Electrical Comm. Engineers of Japan, at Fukuoka, May, 1958. The content of this paper was incorporated in the following paper and is included in the present paper with complete revision. S. MUROGA, "Logical Elements on Majority Decision Principle and Complexity of their Circuit," International Conference on Information Processing at Paris, June, 1959.

(6) S. MUROGA, "Preliminary Study of the Probabilistic Behavior of a Digital Network with Majority Decision Elements," Technical Note RADC-TN-60–146, Rome Air Development Center, August, 1960.

(7) E. GOTO, *et al.*, "Esaki Diode High-Speed Logical Circuits," *Trans. IRE*, Vol. EC-9, No. 1, pp. 25–29 (1960).

(8) M. H. LEWIN, "Negative-Resistance Elements as Digital Computer Elements," Eastern Joint Computer Conference, December, 1959.

(9) E. N. GILBERT, "Lattice Theoretic Properties of Frontal Switching Functions," *J. Math. & Phys.*, Vol. 5, pp. 185–193 (1954).

(10) K. FAN, "On Systems of Linear Inequalities," *Ann. Math. Studies*, No. 38, Princeton, Princeton University Press, 1956, p. 104.

(11) D. G. WILLIS, "Plastic Neurons as Memory Elements," presented at International Conference on Information Processing at Paris, June, 1959.

(12) R. E. GOMORY, "Outline of an Algorithm for Integer Solutions to Linear Programs," *Bull. Am. Math. Soc.*, Vol. 64, No. 5, p. 275 (1958).

(13) A. CHARNES, W. W. COOPER AND A. HENDERSON, "An Introduction to Linear Programming," New York, John Wiley & Sons, Inc., 1953, p. 52.

(14) A. J. GOLDMAN and A. W. TUCKER, "Theory of Linear Programming," Annals of Math. Studies No. 38, p. 62, Corollary 2B, Princeton, Princeton University Press, 1956.

(15) S. MUROGA AND K. TAKASHIMA, "The Parametron Digital Computer MUSASINO-1," *Trans. IRE*, Vol. EC-8, No. 3, pp. 308–316 (1959).

(16) S. MUROGA, "A Program to Discover Functions Representable by a Single Majority Decision Element," (in Japanese) *Electrical Communication Laboratory Technical Journal*, the Nippon Tele. and Tele. Corp., Vol. 8, No. 6, pp. 614–622 (1959).

(17) I. TODA, M. KONDO AND S. MUROGA, "Majority Decision Functions of Six Variables," to be published in the *Electrical Communication Laboratory Technical Journal*, the Nippon Tele. and Tele. Corp.

(18) A part of this table was quoted from the booklet by S. MUROGA, "An Introduction to Automata," (in Japanese), Kyoritusha, October, 1959.

(19) G. BIRKHOFF, "Lattice Theory," New York, Am. Math. Soc., 1948, p. 146.

READER OPINION

SPIRAL STELLAR NEBULAE AND COSMIC GAS JETS

BY

C. E. R. BRUCE [1]

Since I sent you my paper on the above subject (1)[2] I have had my attention called to O. C. Wilson's paper (2) on "The Nuclear and Nebular Spectra of the Planetary Nebula NGC 2392." It will be recalled that this is the nebula which Campbell and Moore (3) took as the type nebula of the group in which the gas movements proved quite inexplicable on any of the many hypotheses they tried. I showed in my Fig. 4 that these movements were at once resolvable into the two diametrically opposed gas jets, initiated by the two electrical discharges which had occurred at some time in the star's atmosphere. The temperature, and hence the gas velocity will vary along each of the two jets in a more or less similar manner, so that any particular level of excitation will be found at corresponding positions in each of the two jets. Thus if a slitless spectrogram of the nebula is obtained it is to be ex-

Photograph by courtesy of Mount Wilson Observatory

pected that there will be two rings, one displaced to the violet, and one to the red of the normal position of the line. This will be seen in the accompanying figure to be confirmed by the slitless image of the line λ3868 [NeIII], taken from Wilson's paper [his Fig. 2b). The brighter of the two rings is the *red* component of the line, indicating that the

[1] Electrical Research Association Laboratory, Leatherhead, Surrey, England.

[2] The boldface numbers in parentheses refer to the references appended to this note.

much greater intrinsic brightness of the further discharge more than counterbalanced the effect of the increased absorption of the light in the star's atmosphere, as compared with that experienced by the light from the nearer of the two jets. Curiously Wilson refers only to the inner "ring," and ignores the existence of the fainter neighbor displaced towards the violet. The outer ring from the nearer jet is evidently too faint to be photographed, and only that associated with the "red" inner ring is seen, itself very much fainter than the corresponding inner ring. The presence of the level of excitation corresponding to λ3868 [NeIII] will depend on both the temperature and the density of the gas. The outer ring is apparently emitted when the jet has almost reached the limits of the star's atmosphere, where presumably the density and the temperature have both decreased, but in such a manner that their combination yields conditions favorable to the existence of the same level of excitation as appears lower down in the star's atmosphere.

However, despite the absence of the violet component of the outer ring, the appearance of the two inner rings confirms the existence of the two jets depicted in Fig. 4 of my paper.

REFERENCES

(1) C. E. R. BRUCE, JOUR. FRANKLIN INST., Vol. 271, p. 1 (1961).
(2) O. C. WILSON, *Ap. J.*, Vol. 108, p. 201 (1948).
(3) W. W. CAMPBELL AND J. H. MOORE, *Pub. Lick Obs.*, Vol. 13, p. 75 (1918).

THE FRANKLIN INSTITUTE LABORATORIES
J. G. RICHARD HECKSCHER, ACTING DIRECTOR

A NOTE ON THE DUCTILITY OF BERYLLIUM SINGLE CRYSTALS ORIENTED FOR BASAL SLIP AND TESTED IN TENSION*

BY

M. HERMAN AND G. E. SPANGLER

This note has been prepared to present experimental evidence which suggests that impurities in beryllium strongly influence the amount of basal plane glide obtained during tensile testing.

Several single crystals of beryllium were prepared by zone melting using the floating zone method. The apparatus used was similar to that previously described for the zone refining of titanium (1)[1] with, however, the addition of a gas re-circulating pump to limit the vapor deposit in the vicinity of the molten zone. Crystals obtained from this process were machined to conical shouldered tensile specimens by electro-chemical methods (2) or a combination of spark-discharge and electro-chemical machining. With the exception of one bend-test specimen (see Fig. 1) the specimens were tensile tested in an Instron

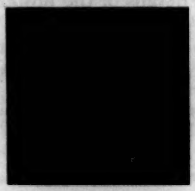

FIG. 1. A tensile specimen (prepared from 8 "pass" material) which was subjected to bending. Orientation, χ_0 approximately 25°; bend axis $[10\bar{1}0]$. Elongation on outer fibers approximately 50 per cent. (Magnification, 1.2×.)

Testing Machine in which the load-extension curves were automatically recorded.

In Table I the history of the specimens and the tensile testing results are presented.

These results, in contrast to previously published work (3) on the tensile deformation of single crystals of beryllium, clearly indicate that the extent of basal glide occurring during tensile testing is not necessarily limited to very small values. Based upon the apparent relationship between the extent of zone melting (number of passes) and the values obtained for the resolved shear stress at yielding, it is suggested that the observed increase in ductility is the result of a decrease in the impurity concentration in the beryllium.

These data are being examined in terms of current theory of the

* This work was supported by the U. S. Navy, Bureau of Naval Weapons.

[1] The boldface numbers in parentheses refer to the references appended to this note.

TABLE I.—*History and Tensile Testing Results of Beryllium Single Crystals.*

Starting Material	Number of Zone Melting Passes	Orientation*	Heat Treatment Prior to Testing	Resolved Shear Stress at Yielding, psi.	Strain at Fracture, % Glide Strain	Operating Glide System
Pechiney (3)	Vacuum Cast	$\chi_0 = 45°$ λ_0 (information not available)	(information not available)	2000 ± 500	~5	Basal Plane
Berylco Vacuum Cast Ingot	2	$\chi_0 = 47°$ $\lambda_0 = 51°$	800°C.–1 hr.	2400	16	Basal Plane
Berylco Vacuum Cast Ingot	5	$\chi_0 = 20°$ $\lambda_0 = 25°$	900°C.–1 hr.	1350	64	Prism Plane
Pechiney	8	$\chi_0 = 48°$ $\lambda_0 = 55°$	None	520	156	Basal Plane

* χ_0 = angle between slip plane and tension axis
λ_0 = angle between slip direction and tension axis

cleavage fracture of single crystals while further experiments are being conducted to examine the details of the proposed impurity effects.

REFERENCES

(1) G. E. SPANGLER, M. HERMAN AND J. L. RUTHERFORD, AFOSR–TN–60–693, 1960.
(2) G. GREETHAM AND A. J. MARTIN, *J. Sci. Instr.*, Vol. 37, p. 160 (May 1960).
(3) G. L. TUER AND A. R. KAUFMANN, "The Metal Beryllium," edited by D. W. White and J. E. Burke, Cleveland, ASM, 1955, p. 367.

THE FRANKLIN INSTITUTE LABORATORIES
J. G. RICHARD HECKSCHER, ACTING DIRECTOR

NOTE ON THE EFFECTIVE CONSTANTS OF COMPOSITE MATERIALS *

BY

Z. HASHIN [1] AND S. SHTRIKMAN [2]

The present work deals with the determination of effective material constants of two-phase isotropic materials, assuming that the macroscopic physical behavior of each of the constituents is analytically known. The specific problem treated is that of the magnetic permeability of two-phase materials. For reasons of mathematical analogy the results obtained also apply to dielectric, electric conduction, heat conduction and diffusion properties of such materials.

The problem is one of long standing (1)[3] and many different results have been obtained. It has been shown by Brown (2) that the effective magnetic permeability is not completely defined by fractional volumes of the constituent phases and that statistical aspects of phase dispersion influence the results. It is the purpose of this work to find the most restrictive upper and lower bounds on the effective constants that can be obtained in terms of phase constants and fractional volumes, assuming isotropy of the composite material.

These bounds are constructed by using the following variational principles (3). Consider a body of volume V and surface S of material of magnetic permeability μ_0, which for simplicity is taken to be constant in space. Let a magnetic potential $\psi_0 = \psi_0(S)$ be prescribed on the boundary. This defines uniquely the field intensity in the body which is denoted by $\vec{H}_0$.

Consider now the same body where part or whole of it has been changed to material with permeability μ, which may vary in space. Define the generalized magnetization by

$$\vec{T} = \vec{B} - \mu_0 \vec{H} \tag{1}$$

where $\vec{B}$ is a magnetic induction and $\vec{H}$ a magnetic field intensity.

* The authors gratefully acknowledge support by the Office of Naval Research and by the Air Force Office of Scientific Research of the Air Research and Development Command.

[1] University of Pennsylvania, Philadelphia, Penna. On leave of absence from the Technion-Israel Institute of Technology, Haifa, Israel.

[2] The Franklin Institute Laboratories for Research and Development, Philadelphia, Penna. On leave of absence from the Weizmann Institute of Science, Rehovot, Israel.

[3] The boldface numbers in parentheses refer to the references appended to this note.

Define

$$\vec{H}' = \vec{H} - \vec{H}_0. \tag{2}$$

Then the volume integral

$$U_T = \frac{1}{8\pi}\int_{(V)} \left(\mu_0 H_0^2 - \frac{T^2}{\mu - \mu_0} + TH' + 2TH_0\right) dV \tag{3}$$

is stationary for

$$\vec{T} = (\mu - \mu_0)\vec{H} \tag{4}$$

where $\vec{H}'$ and $\vec{T}$ satisfy the subsidiary condition

$$\mu_0 \operatorname{div} \vec{H}' + \operatorname{div} \vec{T} = 0. \tag{5}$$

The intensity $\vec{H}'$ must satisfy the boundary condition that ψ', defined by

$$\vec{H}' = -\operatorname{grad} \psi', \tag{6}$$

vanishes on the boundary S.

The expression U_T is an absolute maximum when $\mu > \mu_0$ and an absolute minimum when $\mu < \mu_0$. The stationary value of U_T is equal to the magnetostatic energy stored in the changed body.

A similar variational theorem has been established for a body with prescribed normal component of boundary induction.

Let a two-phase isotropic and quasi-homogeneous material of volume V consist of two phases of volumes V_1 and V_2. The fractional volumes of these are denoted by v_1 and v_2 where

$$v_1 + v_2 = 1. \tag{7}$$

Let μ_1 and μ_2 be the permeabilities of the phases where

$$\mu_2 > \mu_1 \tag{8}$$

and $\vec{T}_1$ and $\vec{T}_2$ the generalized magnetizations in them.

The effective permeability μ^* is defined as follows: Let a potential ψ_0, derived from a uniform $\vec{H}_0$ be applied to the boundary of the composite body. The magnetostatic energy U^* of the body is then taken to be given by

$$U^* = \frac{1}{8\pi}\mu^* H_0^2 V. \tag{9}$$

With reference to Eqs. 1–5 it is now assumed that

$$\begin{aligned} \mu_0 &= \mu_1 \\ \mu &= \mu_2 \end{aligned} \tag{10}$$

$$\begin{aligned} \vec{T}_1 &= 0 \\ \vec{T}_2 &= \vec{T}_0 \text{ constant and parallel to } \vec{H}_0. \end{aligned} \tag{11}$$

Introducing (10) and (11) into (3) and using (5), the following result is obtained

$$U_T = \frac{1}{2}\left(\mu_1 H_0^2 - \frac{T_0^2 v_2}{\mu_2 - \mu_1} - \frac{T_0^2 v_1 v_2}{3\mu_1} + 2T_0 H_0 v_2\right) V \leq U^*. \tag{12}$$

The third term in (12) which corresponds to the third term in (3) requires some clarification. It has already been given by Néel (**4**) and can be derived by solving (5) subject to (6) by Fourier methods (**5**), or by making use (**6**) of the Brown-Morrish (**7**) theorem.

Introducing (9) into (12) a lower bound on μ^* is obtained. The highest lower bound is found when U_T in (12) is maximized. The result is

$$\mu^* \geq \mu_1 + \frac{v_2}{\dfrac{1}{\mu_2 - \mu_1} + \dfrac{v_1}{3\mu_1}} = \mu_1^*. \tag{31}$$

To find an upper bound it is assumed that

$$\begin{aligned} \mu_0 &= \mu_2 \\ \mu &= \mu_1 \end{aligned} \tag{14}$$

$$\begin{aligned} \vec{T}_1 &= \vec{T}_0 \text{ constant and parallel to } H_0 \\ \vec{T}_2 &= 0. \end{aligned} \tag{15}$$

By the same procedure it follows that the lowest bound is given by

$$\mu^* \leq \mu_2 + \frac{v_1}{\dfrac{1}{\mu_1 - \mu_2} + \dfrac{v_2}{3\mu_2}} = \mu_2^*. \tag{16}$$

It should be noted that these bounds are independent of the geometry of structure of the composite material, the only requirement being isotropy.

It can furthermore be shown that these bounds are the most restrictive that can be obtained in terms of the phase constants and fractional volumes. This is done by proving that for certain phase configurations, μ_1^* and μ_2^* given in (13) and (16) are exact expressions for μ^*. For this purpose consider composite spheres each of which consists of an inner

sphere of one phase surrounded by a concentric spherical shell of the second phase, the volume ratio of the two phases being fixed in all spheres. Let a body be completely filled with composite spheres (this is possible if no size restriction is imposed). In this case, when the inner spheres have permeability μ_2 and the shells permeability μ_1

$$\mu_1^* = \mu^*. \tag{17}$$

If the materials are interchanged, then $\mu_2^* = \mu^*$. This follows by a method used by Hashin (8).

The theoretical results given in the present work are verified by experimental investigations of DeLoor (9) for the analogous dielectric problem.

Similar results have been derived by the authors for two-phase elastic materials, formulating and using variational theorems analogous to the ones given here (10).

The authors are indebted to Dr. W. F. Brown, Jr. for valuable ideas generously communicated.

REFERENCES

(1) J. A. REYNOLDS AND J. M. HOUGH, "Formulae for Dielectric Constant of Mixtures," *Proc. Phys. Soc., B*, Vol. 70, p. 769 (1957).

(2) W. F. BROWN, JR., "Solid Mixture Permittivities," *J. Chem. Phys.*, Vol. 23, No. 8, pp. 1514–1517 (1955).

(3) These principles are an extension of a variational theorem formulated by W. F. Brown, Jr. in a private communication.

(4) L. NÉEL, "Le champ coercitif d'une poudre ferromagnetique cubique a grains anisotropes," *C. R. Acad. Sci.*, Vol. 224, pp. 1550–51 (1947).

(5) L. NÉEL, "La loi d'approche en $a:H$," *J. de Physique*, Vol. 9, p. 184 (1948).

(6) S. SHTRIKMAN, "The Magnetic Properties of Small Ferromagnetic Particles," D.Sc. thesis, Technion-Israel Institute of Technology, Haifa, Israel, 1957.

(7) W. F. BROWN, JR. AND A. H. MORRISH, "Effect of a Cavity on a Single-Domain Magnetic Particle," *Phys. Rev.*, Vol. 105, pp. 1198–1201 (1957).

(8) Z. HASHIN, "The Elastic Moduli of Heterogeneous Materials," Nonr 1866(02) TR 9—Div. Engng. Appl. Phys., Harvard University, Cambridge, Mass., 1960.

(9) G. P. DELOOR, "Dielectric Properties of Heterogeneous Mixtures," Thesis, Univ. of Leiden, Leiden, 1956.

(10) Z. HASHIN AND S. SHTRIKMAN, "Note on a Variational Approach to the Theory of Composite Elastic Materials," JOUR. FRANKLIN INST., Vol. 271, p. 335 (1961).

NEW BOOKS IN THE FRANKLIN INSTITUTE LIBRARY

MATHEMATICS

Numerical Methods for Science and Engineering, by Ralph G. Stanton. Prentice-Hall, 1961.

Laplace Transformation, by William Tyrrell Thomson. 2nd ed. Prentice-Hall, 1960.

Complex Variables and the Laplace Transform for Engineers, by Wilbur R. LePage. McGraw-Hill, 1961.

Continuous Geometry, by John Von Neumann. Princeton University Press, 1960.

Tables of Higher Functions, by Jahnke, Emde and Lösch. 6th ed. rev. McGraw-Hill, 1960.

Digital Applications of Magnetic Devices, edited by Albert J. Meyerhoff. Wiley, 1960.

Analogue Computers, by Izrail' Isaevich Eterman. Pergamon, 1960.

PHYSICS

Waves in Layered Media, by Leonid M. Brekhovskikh. Academic, 1960.

Theoretical Physics in the Twentieth Century; A Memorial Volume to Wolfgang Pauli, edited by M. Fierz and V. F. Weisskopf. Interscience, 1960.

The Encyclopedia of Spectroscopy, edited by George L. Clark. Reinhold, 1960.

Radiation Damage in Solids, by Douglas S. Billington and James H. Crawford, Jr. Princeton University Press, 1961.

Variational Principles in Dynamics and Quantum Theory, by Wolfgang Yourgrau and Stanley Mandelstam. 2nd ed. Pitman, 1960.

Symposium on Naval Structural Mechanics. 2nd, Brown University, 1960. Plasticity: Proceedings, edited by E. H. Lee and P. S. Symonds. Pergamon, 1960.

Atomic Theory for Students of Metallurgy, by William Hume-Rothery. 3rd rev. ed. Institute of Metals, 1960.

Elements of Maser Theory, by Arthur A. Vuylsteke. Van Nostrand, 1960.

Hydrodynamics, by Garrett Birkhoff. Rev. ed. Princeton University Press, 1960.

High-Energy Electron Scattering Tables, by Robert Herman and Robert Hofstadter. Stanford University Press, 1960.

Large Elastic Deformations and Non-Linear Continuum Mechanics, by A. E. Green and J. E. Adkins. Clarendon Press, 1960.

An Introduction to Celestial Mechanics, by Theodore E. Sterne. Interscience, 1960.

CHEMISTRY AND CHEMICAL ENGINEERING

The Ring Index, by Austin McDowell Patterson. 2nd ed. American Chemical Society, 1960.

Block and Graft Polymers, by William J. Burlant and Allan S. Hoffman. Reinhold, 1961.

Advances in Fluorine Chemistry. Academic, 1960.

Advances in Analytical Chemistry and Instrumentation. Interscience, 1960.

Introduction to Ceramics, by W. D. Kingery. Wiley, 1960.

MECHANICAL ENGINEERING

Induction Heating: Coil and System Design, by P. G. Simpson. McGraw-Hill, 1960.

Introduction to the Kinematic Geometry of Gear Teeth, by Allan H. Candee. Chilton, 1961.

Design of Worm and Spiral Gears, by Earle Buckingham and Henry H. Ryffel. Industrial Press, 1960.

Die Gasturbine, von J. Kruschik. 2., vollkommen neubearb. und erweiterte Aufl., Springer, 1960.

COMMITTEE ON SCIENCE AND THE ARTS

(Abstract of Proceedings of Stated Meeting held Wednesday, April 12, 1961.)

HALL OF THE COMMITTEE,
PHILADELPHIA, APRIL 12, 1961.

MR. FRANCIS G. TATNALL in the Chair.

The following reports were presented for final action:

No. 3362: Work of Reinhold Rüdenberg.

This report recommended the award of an Elliott Cresson Medal to Reinhold Rüdenberg, of Belmont, Massachusetts, "For his many contributions to the electric power industry for over fifty years as an engineer, inventor, author and teacher, and particularly for his scholarly textbook on transient performance of electric power systems by which present and future engineers may attain a better understanding of the problems of their profession."

No. 3384: Negative Impedance Booster.

This report recommended the award of an Edward Longstreth Medal to Josiah L. Merrill, Jr., of Murray Hill, New Jersey, "For his analysis of problems associated with losses and distortion in telephone transmission systems; for his synthesis of the negative-impedance booster; and particularly for his successful development of a negative-impedance booster system that has proved to be highly effective and economical."

D. S. FAHRNEY,
Secretary to Committee

BOOK REVIEWS

SENSORY DEPRIVATION, edited by Philip Solomon, Philip E. Kubzansky, P. Herbert Leiderman, Jack H. Mendelson, Richard Trumbull and Donald Wexler. 262 pages, diagrams, 5½ × 8½ in. Cambridge, Harvard University Press, 1961. Price, $5.75.

Man is always searching for answers to the mysteries about his kind. And especially about his central function, which puts him at the top of the present phylogenetic tree. When the time comes that we can roadmap the anatomy of the brain, spinal cord, and peripheral nerves, down to the last neuron, maybe we can reach some certainties. With our present techniques and human limitations, this is too great a task.

The physiology of single nerve function and transmission has been examined and understood much more thoroughly than the anatomy or physiology of the central nervous system as a unit. In 1957, the mechanism of transmission of a single impulse from the endoplasmic reticulum to the cell membrane was demonstrated by the British. Only in recent months have the neurohormones and neuroreceptors involved in transmission between cell membranes been elucidated. Our understanding of central function lags far behind this.

The problem of collecting objective data from a system so beset with undefinable and unknown variables is truly monumental.

In the last thirty five years techniques have been devised for obtaining data along many parameters of central function. Out of these developed several useful hypotheses. But the limiting problem is always the same: in the infinity of variables, there is not one, or preferably several hundred thousand, which can be adequately controlled and interpreted. Were this possible, we would have a chance to develop a starting point. Of course, there are anatomical and ethical prohibitions to objective measurement, too.

The real problem in initiating a systematic study of central function is breaking into the circle of variables with one constant. Although the literature contains many reproducible observations, this goal is probably some years in the future. Studies of sensory deprivation may shorten this interval.

Stemming from the investigations of "brainwashing" during the Korean War, and carried forward with mushrooming impetus in anticipation of manned space travel, the concept of depriving a human of the common stimuli to which he is constantly exposed (light, taste, smell, sound, light touch, deep pressure, heat, cold, pain, the kinesthetic senses, etc.), and studying the responses to this new stimulus deficit, has provided an entire new area for behavioral investigation. The field is only several years old, and so far, not much common ground has appeared among investigators.

The purpose of this symposium in 1958 was to bring some unity and standardization to the area. As well as offering some excellent criteria toward this end, a cross section of revealing behavioral experiments are unfolded, adding to our understanding of why you and I do what we do.

From the data many useful conclusions are evident. New versions of the stimulus-response hypothesis are offered, and the new concept that the reticular substance of the brain stem has cognitive and integrative functions is suggested.

Though this is all valuable in itself, the measure of greatness of this investigative tool is that it may provide the break in the circle —the first controlled variable—not by adding measurable stimuli to the system, but by removing them from it—so that we may begin the systematic study of central function.

This symposium is well balanced enough to be considered a standard textbook of psychology at the graduate level. The material is of interest to psychiatrists, and the greater meaning of it will be challenging to all biological philosophers.

CHARLES W. ANDERSON
Jefferson Medical College

QUANTUM THEORY OF ATOMIC STRUCTURE, VOL. II, by John C. Slater. 439 pages, 6 × 9 in. New York, McGraw-Hill Book Co., Inc., 1960. Price, $13.00.

While this second volume of *Quantum Theory of Atomic Structure* (Chapters 17 to 26) is a natural sequence of the first volume, the level of treatment is considerably higher and the subject matter much more specialized. The Hartree self-consistent field method was dealt with in the first volume. The second volume begins with the Hartree-Fock method. The treatment of this method is thorough and beautiful, and many recent contributions of Slater and others are included, in particular an illuminating discussion of the "exchange potential." After a discussion of correlation problems, the application of group theory to atomic structure is exposed clearly but rather briefly (30 pages). The quantum mechanical tools thus sharpened are used to study more rigorously the more complex non-relativistic aspects of atomic spectra, namely, Multiplet theory and Racah's coefficients. Dirac's relativistic theory of the electron follows, and its physical implications are fully discussed. The magnetic behavior of complex atoms, and transition probabilities of radiation are analyzed. The last chapter is devoted to hyperfine structure (magnetic dipole, and electric quadrupole H.F. structure of one electron atom).

The book is carefully and pedagogically written. It is an up-to-date synthesis of the subject matter considered, since so many of the contributions of the last twenty years have been interpreted for the first time in a single coherent picture. Some of the appendices and tables will prove to be considerably helpful but one of the most useful features of the book is probably the excellent and thorough 54-page bibliography extending over the 1920–1960 period, with the full titles of the papers given. This bibliography is invaluable to all physicists engaged in this field of research. The main reservation in the opinion of this reader is that group theory was treated too briefly, maybe half-heartedly and that the elegance of group-theoretical methods has not been fully exploited.

In summary, Slater's two volumes on atomic structure are almost sure to become classical. While the first volume is quite suitable as a textbook, the second will be a basic reference book. It is quite likely that many readers of Slater's books are anxious (like the present writer) to see the promised volumes III, IV, etc . . . . dealing with molecules and solids.

HENRI AMAR
Temple University

PRINCIPLES OF METEORITICS, by E. L. Krinov. 535 pages, illustrations, 5½ × 8½ in. Oxford, Pergamon Press, Ltd., 1960. Price, 70s net.

The advent of the space age has given new meaning to meteoritics. Today, by means of satellites and rocket probes, the scientist is consistently sampling space to uncover its content of particulate matter.

In the past we were passive observers who could only study comets when they came into the sky, who could only see meteors when they flashed across our sky and who could examine meteroites only when they fell to Earth. As a result of our passive role the volume of information coming to us was strictly limited; yet even with these limitations there have been many in the field who have contributed significantly to our knowledge. Names like Olivier, Wylie, Ninninger and others come to mind when we discuss things that fall from the sky. However, these are American astronomers and researchers who have devoted their lives to this study. Others in other lands have also contributed significantly to this science. As Harrison Brown in his preface writes: "It [this book] is admittedly nearly as deficient in reference to English literature on the subject as our works are deficient on reference to Russian literature."

One of the strongest teams has been the Russians who have diligently pursued this science for a long time. Names like Kulik, Levin, Fesenkov and others in the Soviet Union are well known for their contributions. The author, like the others, is an authority in this field.

Principles of Meteoritics is a well written, impressive treatise on meteoritics. The author has covered the various facets of this science in a lucid and brilliant fashion.

The book begins in the traditional manner with the historical background of meteorites and then covers the motions and fall of these objects from the sky. In this section many

of the questions posed to astronomers such as the speed of meteors, sounds that accompany them, brightness, sizes, etc., are answered. The mechanism involved in the burning out of a meteor is also covered though better explanations have appeared in English texts. Interesting tables going back to 1798 are included indicating the fall of meteorites which have struck buildings or people. The most prominent of the meteor craters on the Earth are also described and discussed. In the case of the Tunguska meteorite the author indicates that the latest theory is that this meteorite was really the nucleus or head of a comet which struck the Earth to account for the odd effects felt on impact.

One of the most significant sections of this book is on the morphology of meteorites. These studies of the dimension, form, surface relief, structure of the fusion crust and of the inner structure is most important in our understanding of these objects. The author has performed a magnificent bit of writing in this section to account for the variations we find in meteorites.

For the first time this reviewer has found an explanation as to why there are the variations in the widmanstatten figures in meteorites. The author also gives a quick method of checking the authenticity of a meteorite by its nickel content. He says that while the nickel content of terrestrial iron does not exceed 2–3 per cent, no meteorite has been found with a nickel content of less than 4%. Some go as high as 30 per cent or even more.

Following two sections on the chemistry of meteorites and the mineralogy of meteorites there is a section on the physical properties of meteorites. Here the author explores the specific gravity, melting points, magnetic properties and optical properties of meteorites.

In two appendices are listed the meteorites which are in the USSR and a catalogue of rare meteorites.

There is one omission of major significance. The author has not included in his book results, current findings, on particulate matter in space which falls to the Earth. These results are rather definitive as they come from rocket probes and from satellites. Instead of the 6.6 tons falling to the Earth in a 24-hr. period, the latest results—which were available to the author—indicate that between 1000 and 10,000 tons fall to the Earth in this period.

Principles of Meteoritics is a brilliant work and promises to be a ranking reference book on meteoritics for a long time.

I. M. LEVITT
The Fels Planetarium

MECHANICS OF SOLIDS AND FLUIDS, by Robert R. Long. 156 pages, diagrams, 6 × 9 in. Englewood Cliffs, Prentice-Hall, Inc., 1961. Price: $9.00 (trade); $6.75 (classroom).

Written for the instruction of undergraduate engineering students, this introduction to the analytical mechanics of continuous media makes consistent and effective use of vector and tensor methods, to which the introductory chapter is devoted. Their application to the solution of problems in stress, deformation and flow is illustrated in the two succeeding chapters. In accordance with the author's intent to write an engineering book rather than one on mathematical physics, the fourth chapter discusses the mechanical properties of materials and points out the extent to which most of them differ from the two convenient abstractions, isotropic elastic solids and Newtonian fluids.

The tightly yet clearly written fifth chapter derives the equations of elastic solids with the use of tensor notation in only thirteen pages, and provides an exhibition of its power. Chapter 6 illustrates the application of these equations to a selected group of problems in linear elasticity. Chapters 7 and 8 deal, respectively, with the equations of fluid mechanics and their application to specific problems in both statics and dynamics.

The author has admirably succeeded in writing an introductory text on the mechanics of continua, which is rigorous and complete enough to serve as a foundation for the study of more extensive and specialized works. Yet the style is both lucid and informal, and clearly displays the author's skill in teaching a fundamental engineering subject. There are many thought-provoking examples, and the references are not merely cited, but evaluated in their specific context, with the material discussed in each chapter.

KLAUS CAPPELL
The Franklin Institute Laboratories

BOOK NOTES

LINEAR GRAPHS AND ELECTRICAL NETWORKS, by Sundaram Seshu and Myril B. Reed. 315 pages, diagrams, 6 × 9 in. Reading (Mass.), Addison-Wesley Publishing Co., Inc., 1961. Price, $9.75.

Introducing the theory of linear graphs to electrical engineers and demonstrating the power of these methods in solving network problems, are the aims of this text.

The first five chapters are devoted to the basic theory of graphs in relation to their engineering applications. Considerable space is devoted to the matrices of a graph, and the relevant concepts are also treated in great detail. The last five chapters, which comprise the majority of the book, treat applications of graph theory. These deal with electric network theory, switching theory, communication networks, signal flow graphs, and logic.

The book establishes a solid mathematical foundation for electric network theory, particularly those formulations which are applicable to digital computers, and is based upon the results of several years' research and teaching by the co-authors.

COBALT: CHEMISTRY, METALLURGY AND USES, edited by Roland S. Young. 424 pages, diagrams, 6 × 9 in. New York, Reinhold Publishing Corp., 1961. Price, $15.00.

The chemist who has occasion to work with cobalt and requires accurate, up-to-date and complete information, and the chemist or metallurgist who wishes to browse through this field in the hope of adding to his general knowledge or acquiring ideas which may apply to his own sphere of work, are the ones to whom this book is directed.

The latest information on the production and consumption of cobalt in all its forms, and a complete survey of the many new industrial applications of cobalt that have been developed in the past decade, are found in this text. The properties and compounds of this increasingly important metal are discussed by leading authorities in the field, in addition to the occurrence and extractive metallurgy of cobalt.

MAGNETIC TAPE INSTRUMENTATION, by Gomer L. Davies. 263 pages, diagrams, 6 × 9 in. New York, McGraw-Hill Book Co., Inc., 1961. Price, $8.50.

This guide to the uses of magnetic tape reproduction includes information on techniques and equipment for reproduction of data (not voice and music), such as electrical signals or physical phenomena which can be represented by electrical signals. It is particularly useful as a new technique for scientists and engineers, for it requires no processing between recording and reproduction. The book is aimed at users of magnetic tape equipment, to give them a broader understanding of this new tool. It is not a designer's manual, but a user's one, since applications are stressed.

SEPARATION PROCESSES IN PRACTICE, edited by Robert F. Chapman. 209 pages, diagrams, 5 × 7¼ in. New York, Reinhold Publishing Corp., 1961. Price, $4.95.

The latest information on both theory and practice of the various separation processes, based on fundamental processes, is presented in this small text. The newest theories are examined and presented in the first two chapters along with a review of the fundamentals. Typical examples of both the oldest and the newest techniques employed in the separation process are discussed in the remaining chapters. Described also is some of the latest equipment now in use in the field. This work is based on a symposium, "Experiment in Industry" which was jointly sponsored by the Philadelphia-Wilmington Section of the American Institute of Chemical Engineers and the University of Pennsylvania. The

engineer who desires to brush up on fundamentals, broaden his practical knowledge of the field, and learn of new developments will find this book of valuable assistance.

MODERN FLIGHT DYNAMICS, by W. Richard Kolk. 288 pages, diagrams, 6 × 9 in. Englewood Cliffs (N. J.), Prentice-Hall, Inc., 1961. Price: $10.00 (trade); $7.50 (text).

This volume represents the first in the forthcoming Space Technology Series. Its treatment is as a text and reference in the analysis of flight-vehicle design for aeronautical and astronautical engineers or for anyone engaged in the design or fabrication of airplanes and missiles. Chapter One begins with Newton's Second Law and Lagrange's Equation as alternative ways of formulating the equations of motion. The author then introduces in the second chapter direction cosines and Eulerian angles for the purpose of orientation reference. These two introductory sections are then used as tools with which the dynamics of flight vehicles are analyzed. The author also, in dealing with aerodynamic forces and moments, has not reduced equations of motion to nondimensional status for purposes of generality as do most books on aeronautics, but has treated all motion as *true* and seconds as constant quantities—not varying with speed, altitude, and wing loading. For purposes of clarity and east of continuity, the text retains the "body-fixed" axis system throughout, rather than using the "velocity-fixed" system in some problems. Important features include a complete analysis of the roll-coupling phenomenon, the detailed description of manual and powered control systems for aircraft in Chapters 9 and 10, and in Chapter 12, the approximate solution for the gyroscopic motion of a spin-stabilized rocket. Many illustrative examples are used to emphasize the analyses, and in addition numerous line drawings, graphs, and tables have been included.

TRANSISTOR CIRCUIT ANALYSIS, by Maurice V. Joyce and Kenneth V. Clarke. 461 pages, diagrams, 6 × 9 in. Reading (Mass.), Addison-Wesley Publishing Co., Inc., 1961. Price, $10.75.

The basic methods of analysis involved in the understanding and design of junction transistor circuitry are presented in this text. The transistor models used are restricted to a small number that are all easily interrelated. Initially, the simple low-frequency models are presented and are related to the underlying physical processes that occur in the device. Then after the reader has completely familiarized himself with these simple models the complications, such as high-frequency effects, noise, or saturation effects, are introduced. Special features of this text include a unified approach to transistor biasing problems, use of the current peak of the tunnel diode as a reference device, and an outline of the use of a minority charge density model in transistor switching and saturation calculations. New material is also included on emitter-follower transient responser, video amplifier with maximum flatness and maximum gain bandwidth products, and distributed amplifiers. The book can best be used as a reference for the circuit designer and as a senior-graduate level text.

ADVANCED DESIGN IN STRUCTURAL STEEL, by John E. Lothers. 583 pages, diagrams, 6 × 9 in. Englewood Cliffs (N. J.), Prentice-Hall, Inc., 1961. Price: $15.00 (trade); $11.25 (text).

Practical and relatively simple solutions to the numerous complex and critical problems faced by the modern steel design engineer are provided in this volume. Up-to-the-minute data on continuous and rigid frame steel buildings and statically indeterminate trusses have been included. Wind and earthquake stresses, floor and stair design, and steel frame connections for multi-story buildings are also appraised. The author has included such topics as light gage structural steel sections, torsional buckling, allowable web sheer and tension members. Important information is also given regarding Lamella roofs and one-story rigid frame bents. Numerous illustrations to clarify factual material, and valuable bibliographies with sources for methods of advanced analysis have also been included. Meeting all the requirements and demands of specialists in structural engineering, this book will assist industrial concerns, technical institutes, and government agencies.

PHYSICS OF SEMICONDUCTORS, by A. F. Ioffe. 436 pages, diagrams, 6 × 9 in. New York, Academic Press Inc., 1961. Price, $12.50.

Translated from the Russian work of 1957, the English version of *Physics of Semiconductors* has been revised and supplemented by the author, who is Director of the Institute of Semiconductors for the U.S.S.R. Seven chapters cover thoroughly solid electrolytes, metals, electronic semiconductors, quantum theory of semiconductors, properties of the crystal lattice, methods of measuring the basic properties of semiconductors, and semiconducting materials. A final chapter gives the author's conclusions briefly and clearly.

BOOLEAN ALGEBRA AND ITS APPLICATIONS, by J. Eldon Whitesitt. 182 pages, diagrams, 6 × 9 in. Reading (Mass.), Addison-Wesley Publishing Co., Inc., 1961. Price, $6.75.

Designed as an introduction to Boolean algebra for readers with a limited formal mathematical background, this text can be used in a one-semester course. It is also suitable as background material for specialized courses in the three major applications of Boolean algebra—symbolic logic, probability theory, and circuit design. The seven chapters cover algebra of sets, the disjunctive and conjunctive forms of Boolean algebra, symbolic logic and the algebra of propositions, switching algebra, relay circuits and control problems, circuits for arithmetic computation, and probability in finite sample spaces.

INTERFEROMETRY. National Physical Laboratory Symposium No. 11. 471 pages, illustrations, 6¾ × 9¼ in. London, Her Majesty's Stationery Office, 1960. Price, $5.60 postpaid. (U. S. agent, British Information Services.)

Interferometry is the proceedings of a symposium held at the National Physical Laboratory in June, 1959. Nineteen papers and the resulting discussions comprise the text. Since all the papers were circulated in advance, the discussions form a valuable part of the proceedings. Primarily for physicists, the symposium covered recent developments in interferometry, particularly as related to metrology, optics, radio and radiation.

PERCHLORATES: THEIR PROPERTIES, MANUFACTURE AND USES, edited by Joseph C. Schumacher. ACS Monograph No. 146. 257 pages, diagrams, 6 × 9 in. New York, Reinhold Publishing Corp., 1960. Price, $8.75.

All the current knowledge of the technology of perchlorates and perchloric acid has been brought together in this new ACS monograph, by leading authorities in the field. They have traced the commercial development of the perchlorate industry on a world-wide scale and reviewed the properties and manufacture of the compounds, discussing their safe handling and examining their various uses. An entire chapter is devoted to the application of perchlorates in the explosive and propellant industries. This material is particularly timely due to the growing interest in the compounds as oxidizers in propellants for the military space program. This book, aside from its immediate value as an accurate reference for present day problems encountered in the manufacture and use of perchlorates and perchloric acid, provides the background for future research and development in this expanding field.

LINEAR CIRCUITS, by R. E. Scott. PART I, TIME-DOMAIN ANALYSIS, 510 pages, diagrams, 6 × 9 in. Price, $6.75. PART II, FREQUENCY-DOMAIN ANALYSIS, 418 pages, diagrams, 6 × 9 in. Price, $6.75. Reading (Mass.), Addison-Wesley Publishing Co., Inc., 1960.

This introduction to linear circuits for students of electrical engineering presents the methods of linear circuits as a solid foundation upon which the complete structure of electrical engineering can be built. The author assumes in this work that the student has had an introductory course in physics and is studying calculus concurrently. Part I, containing fourteen chapters, treats the basic theory, and topics such as singularity functions, response functions in the time domain (that is, step and impulse responses), differential equations, and convolution theory. Part II, with nine chapters, treats the network system functions in the frequency-domain (that is, responses to steady-state sinusoids), and Laplace transforms and power density spectra for more general input functions.

The text also covers topics seldom found in books at this level, such as signal-flow graphs, relaxation methods, dummy variables, response functions in the time and frequency domain, and s-plane plots.

INTRODUCTION TO THE THEORY OF IONIZED GASES, by J. L. Delcroix. Translated by Melville Clark, Jr. and David J. BenDaniel. 149 pages, diagrams, $5\frac{1}{4} \times 8$ in. New York, Interscience Publishers, Inc., 1960. Price: $4.50 (cloth); $2.50 (paper).

Translated from the 1959 French edition, this text approaches the theory of ionized gases through a microscopic description of the state of the gas. This complements Spitzer's well known work, using macroscopic equations. The material has been tested in courses given by the author at the Faculté des Sciences (Paris) and the Institut des Sciences et Techniques Nucléaires (Saclay). Twelve chapters deal with such topics as elastic collisions, the kinetic theory of gases, the theory of free electron gases, weakly ionized gases, and plasmas. Four short appendices give some of the mathematical derivations. This is one of the few available works in the new field of plasma dynamics.

52 FOURIER TRANSFORMS, by R. R. Goldberg. 76 pages, $5\frac{1}{2} \times 8\frac{1}{2}$ in. New York, Cambridge University Press, 1961. Price, $3.75.

This study provides a background in classical Fourier transform theory, in a form which lends itself to an understanding of abstract harmonic analysis. The introductory chapter contains all the theorems on integration which are used in the remaining four chapters dealing with: the Fourier transform on L^1; the Fourier transform on L^2; generalizations of Wiener's theorem; and Bochner's theorem. An appendix points out the relationship to an arbitrary locally compact abelian group. The reader is assumed to have a knowledge of Lebesgue and Riemann-Stieltjes integration.

INDUSTRIAL ORGANIC NITROGEN COMPOUNDS, by Melvin J. Astle. ACS Monograph No. 150. 392 pages, diagrams, 6×9 in. New York, Reinhold Publishing Corp., 1961. Price, $14.00.

This monograph summarizes the chemistry of organic nitrogen compounds, with chapters on: the aliphatic amines; arylamines; heterocyclic amines; hydrazine, azo compounds, diazonium salts and oximes; nitrides, amides and amino acids; isocyanates, ureas, thioureas; aromatic nitro and nitroso compounds; and aliphatic nitro compounds. Patent literature receives a great deal of attention, so that industrial as well as theoretical developments are included. The simple heterocyclic nitrogen ring is discussed in detail, as background for an understanding of the nitrogen compounds. Some topics, such as amino acids and proteins, alkaloids, and heterocyclic nitrogen compounds, do not receive as much attention as the others, since they have been covered thoroughly in other works.

HYPERSTATIC STRUCTURES. VOL. 2: WORKED EXAMPLES AND EXAMPLES FOR SOLUTION, by J. A. L. Matheson and A. J. Francis. 282 pages, diagrams $6 \times 9\frac{1}{2}$ in. New York, Academic Press Inc., 1960. Price, $11.00.

This work has been published in two complementary volumes, the second volume containing worked examples and problems applying the principles and methods discussed in Vol. 1. The subject of hyperstatic structures (a structure is said to be "statically indeterminate" or "hyperstatic" when it possesses more members or is supported by more reactive restraints than are strictly necessary for stability) is treated as an organic and coherent whole by the authors. The relationship between the various energy theorems is demonstrated by a consideration of non-linear structures; then the theorems and methods applying only to linear structures are treated in their logical order, with special attention being paid to the broad distinction between the "equilibrium" and "compatibility" approaches. As an introduction to the analysis of structures on the digital computer, a chapter on matrix methods is included.

ELEMENTS OF STATISTICAL INFERENCE, by David V. Huntsberger. 291 pages, diagrams, $6\frac{1}{2} \times 9\frac{1}{2}$ in. Boston, Allyn and Bacon, Inc., 1961. Price, $7.25.

This book could serve as a text for a general course for those who want to know "what

statistics is all about." The techniques and methods which are presented are those which are generally applicable, and methods which are of interest in but a few areas of application have not been included. Some of the examples and problems set forth in the book are artificial, some have been abridged from real sets of data, and others are the unabridged results of samples and experiments.

A knowledge of arithmetic and some exposure to algebra are sufficient for the reader, since the mathematical prerequisites have been kept to a minimum. Those who want to be familiar with at least the language and fundamental concepts of statistics in order to achieve additional competence in their chosen fields ,would find this work beneficial. In summary, the author intended this book as an introduction to some of the basic concepts and techniques of statistical inference that play an important role in the solution of problems that arise in many spheres of human activity, and he hopes to give the reader some understanding of the statistical approach to problems faced by the administrator, the research worker, and other members of our increasingly complex society.

NUCLEAR ENGINEERING, by Gilbert Cahen and Pierre Treille, translated by Gilbert B. Melese. 394 pages, illustrations, 6½ × 8¾ in. Boston, Allyn and Bacon, Inc., 1961. Price, $8.50.

Published originally as the highly successful "Précis d'Énergie Nucléaire" by the two authors listed above, two of France's leading nuclear engineers, this book has been translated and brought up-to-date by Dr. Melese to include the latest available information plus pertinent data from the 1958 Geneva International Conference.

Those elements of modern physics which are significant in nuclear engineering are discussed in the first two chapters, followed in succeeding chapters by discussion on nuclear reactor theory, nuclear reactor materials, and French, British and Russian reactors. The book offers conversion tables, nuclear data tables, charts and graphs, and concludes with an examination of the hazards as well as the beneficial uses of radiation and includes a discussion of radiation shielding. The original French version of this work has been reviewed as "The best introduction to the immense domain of nuclear energy," and has been widely used in Europe with great success.

PUBLICATIONS RECEIVED

SCIENTIFIC RUSSIAN, by James W. Perry. Second edition, 565 pages, 6 × 9 in. New York, Interscience Publishers, Inc., 1961. Price, $9.50.

CHARACTERISATION OF ORGANIC COMPOUNDS, by F. Wild. Second edition, 306 pages, 5½ × 8½ in. New York, Cambridge University Press, 1961. Price, $2.95 (paper).

ADVANCED CALCULUS, by David V. Widder. Second edition, 520 pages, 6 × 9 in. Englewood Cliffs (N. J.), Prentice-Hall, Inc., 1961. Price: $12.00 (trade); $9.00 (text).

THERMODYNAMICS, by G. N. Lewis and M. Randall, revised by Kenneth S. Pitzer and Leo Brewer. 723 pages, diagrams, 6 × 9 in. New York, McGraw-Hill Book Co., Inc., 1961. Price, $12.50.

DICTIONARY OF MECHANICAL ENGINEERING, by Alfred Del Vecchio. 316 pages, 6 × 9¼ in. New York, Philosophical Library, 1961. Price, $6.00.

DAILY RIVER STAGES, VOL. LIV, 1958. 174 pages, 8¼ × 10¾ in. Washington, U. S. Dept. of Commerce, 1960. For sale by the Superintendent of Documents, U. S. Gov't Printing Office, $1.25 (paper).

NATIONAL BUREAU OF STANDARDS NEWS

X-RAY BARRIER DETERMINATIONS

Measurements made by the National Bureau of Standards, in a program sponsored by the Atomic Energy Commission, indicate that present methods of determining barrier requirements for X-ray installations may result in walls having protective barriers that are several half-value layers thicker than actually required.[1] This study, conducted by S. W. Smith and J. R. Brooks, with the cooperation of over 20 hospitals and clinics, may lead to savings in the cost of many future installations.

The usual procedure for the calculation of required barrier thickness is to assume near-maximum values for the various factors affecting the exposure of the barriers if more precise knowledge of them is not known. These factors include such variables as workload, fraction of time the radiation is directed toward the barrier, effect of coning, attenuation of the patient, and others. In actual practice, one or more of these factors may be considerably less than the maximum values assumed for busy installations and the barrier thickness required for personnel protection may thus be less than that indicated by the calculations based on these assumptions.

In this study, the amount of radiation reaching various regions of the X-ray rooms was determined through the use of pocket-type ionization chambers. As many as 70 of these instruments were attached to selected positions on the ceiling and walls of each room. Readings were taken once or twice a week, depending on the exposure, for a period of from one to two months. The penetrating quality of the radiation was determined from attenuation curves obtained from readings of instruments that were surrounded with varying thicknesses of aluminum or lead.

The work load in milliampere-minutes of each installation was automatically determined by a charge-integrating device connected to the X-ray control. From the data obtained it seems reasonable to assume, when no specific information is available, a workload of 1000 ma-min/week for a busy diagnostic X-ray room.

The amounts of lead shielding required to reduce exposure to the desired level were calculated from the collected data for those rooms having maximum exposure and use. These thicknesses were generally

[1] For further technical details see, "Measurement of Radiation Exposure at the Walls of Medical X-Ray Rooms for Determining Protective Barrier Requirements," by S. W. Smith and J. R. Brooks, *Atompraxis*, March, 1960.

considerably lower than values arrived at using the usual procedure, indicating substantial safety factors.

The results of this study, combined with data from other experiments, may eventually result in a reduction in the amount of shielding specified for X-ray rooms by use of more realistic values for the various factors used in computing the barriers.

YOU CAN ADVANCE SCIENCE EDUCATION

Today, more than ever before in its 137-year history, there is vital need for The Franklin Institute to effectively promote education in science and technology. It is imperative that we meet this challenge by providing adequate educational opportunities in these fields. This requires vision, objective planning, and money. We have more than enough of the first two requisites, but far too little of the third.

Our programs are aimed at professional scientists and industry, as well as the lay public and young people seeking inspiration and guidance in choosing a career. The Institute's educational programs are impressive. They begin with students in the early grades of our elementary schools and continue throughout an individual's professional or industrial life. With more funds at our disposal, the scope and vigor of these activities could be greatly increased and increasingly effective.

The Franklin Institute is not richly endowed. It is a non-profit organization, depending for encouragement and support on an understanding public. Capable and conservative management assures wise administration of all funds.

Your gift or bequest, large or small, will be deeply appreciated and will be used effectively to broaden the Institute's educational usefulness. There is a warm satisfaction in giving financial support to an organization that has pioneered in, and is dedicated to, the advancement of science and technology.

When property is transferred, title should be in the name of The Franklin Institute of the State of Pennsylvania for the Promotion of the Mechanic Arts.

The Secretary of The Franklin Institute will gladly furnish you with additional information. Write to him at The Franklin Institute, Benjamin Franklin Parkway at Twentieth Street, Philadelphia 3, Pennsylvania.

CURRENT TOPICS

"Phantom Holes" in Computer Tape.—A problem caused by oil spots which, in effect, became "phantom holes" in paper tape for instructing computers, and led to faulty data-processing results, has been solved by the Radio Corporation of America.

Solution of this problem—common throughout the computer industry—has been achieved by RCA through development of a high-speed tape reader that knows a genuine perforation when it sees one.

Much of the paper tape employed in present-day electronic data processing is oil-impregnated and small areas of oil concentration sometimes occur. Oil spots also may be caused by lubricants in connection with high-speed operation of data-processing equipment.

Photoelectric scanning principles and improved circuitry have been incorporated in the design of RCA's Model 322 Tape Reader, enabling it to ignore everything but true perforations.

In testing the selectivity of the new tape reader, RCA engineers deliberately sprinkled perforated paper tapes with oil and then ran them through the tape reader. The results were uniformly successful, he said.

There are two models of the new RCA paper tape equipment—one serves as a tape reader only, while the other can read as well as punch the data-signifying holes. The reader can scan perforated tape at rates up to 1,000 data characters per second, handling 5, 6 or 7-level formats. Accuracy in the reading operation is assured by a parity check.

Resembling an undersized secretary desk, the Model 322 Paper Tape Reader/Puncher stands 60 in. high and measures only 34 in. by 24 in. at the base.

New Atomic Clock.—An atomic clock, so accurate that its maximum error would not exceed one second in 1271 years, has been developed for the Air Force. The airborne device which uses atomic energy as a power source is technically referred to as an Airborne Atomic Frequency Standard. It should become operational for use in Air Force missiles and aircraft next year.

The 62½-lb. airborne atomic frequency standard will replace the numerous crystal oscillators, used as frequency or time standards, now required to calibrate airborne communications, navigation, guidance, fire control, computers and timing devices.

First of its kind to be developed, the unit has been simplified to the extent that only a simple on-off switch is required to operate it. Present oscillators require frequency calibration, often just before takeoff or during flight. The new atomic standard does not require calibration, acting as its own primary standard. Operator of the unit needs no special schooling such as is required on present laboratory units.

The small clock not only will be airborne, but can also replace the present-day atomic laboratory units weighing over 600 pounds. No radiation hazard is involved in its operation.

The airborne atomic frequency standard is accurate to one part in 10 billion and is stable to one part in 25 billion.

Combining the latest techniques in solid state devices, the atomic frequency standard was developed to airborne specifications by the National Company, Inc., Malden, Mass., under the guidance of the Communication Laboratory, Wright Air Development Division, Air Research and Development Command, Wright-Patterson AFB, Dayton, Ohio.

Telemetry System for Space.—General Electric's Missile and Space Vehicle Department in Philadelphia has completed an experimental version of a new telemetry system that is expected to be able to beam signals through space about three times as far as any telemetry system yet flown, including Pioneer V, United States solar satellite which last year exceeded all records by sending signals to Earth from millions of miles out. The telemetry system is one of a family of communication systems called Synchrolink that will transmit data the same distance as systems now in use with about one-tenth the power requirements. The estimated ranges are based on the inherent efficiencies of the system, which is described as "the most efficient practically realizable with today's technical know-how."

Synchrolink is a digital system, which gives it a high degree of data accuracy. It is classified as Pulse Code Modulation with Phase Shift Keying (PCM/PS). An optional feature is an error correction code which can be used to correct mistakes caused by noise interference.

Other reasons Synchrolink is attractive are:

1. Relatively minor modifications would be needed to convert existing ground stations to receive Synchrolink signals.

2. Transistors are used in their most reliable mode: as switches.

3. It is flexible: channels and combinations of channels could be altered to accommodate measurement changes, even in flight.

4. Repeated conversion, or transmission of data, or long storage, causes no loss in accuracy.

In a typical application, information such as the space vehicle performance and temperature environment is gathered by sensors located throughout the vehicle. Samples of information direct from the sensors, or via pre-transmission analyzers, are coded into a series of positive or negative pulses. The Synchrolink system then beams a radio frequency signal that is switched 180 degrees in accordance with a simple coded pulse series representing the space flight data. The resulting double side band-suppressed carrier radio frequency signal is received by a synchronous receiver followed by a correlation detector. Use of the threshold-free synchronous receiver permits detection of the signals at extremely low signal-to-noise ratios, about -10 db.

The detected output is recorded on tape and run through decoding and processing equipment most of which is standard.

Components contributing to the unique properties of the Synchrolink system are: Error Correction Encoder, Modulator for Phase Shift Keying, Synchronous Receiver, Correlation Detector, and Error Correction Decoder.

Classroom Alpha Particle Counter.—A new Open Air Alpha Counter for demonstrations in the high school and junior college classroom has been announced by Central Scientific, Chicago.

The apparatus is able to detect alpha particles as they pass through the instrument causing a dielectric

breakdown of air under a high potential gradient. This breakdown, as indicated by a small spark, occurs whenever a particle ionizes a path in the air between the Counter's nickel-chromium grid and detector plate. These two are maintained at about 5000-volt d-c. potential difference by an accessory power supply. The source of alpha radiation (radium salts) is mounted on an adjustable nylon rod that can be moved back and forth in relation to the counter grid. By varying the d.-c. potential or the position of the source, the rate of counts can be closely controlled. When operated with a scaler, an accurate count of particles can be made over a pre-determined time interval.

A clear plastic cover protects the radioactive source and counting grid from student contact and dust during storage. The Counter is mounted on a vertical rod in a stable Harrington base. The instrument requires little maintenance other than periodic cleaning of the grid wires with carbon tetrachloride.

The counter operates at a maximum potential of 5000 volts d.c. at a maximum of 1 milliampere. A variable 5000-volt d-c. source is also available from Cenco.

Moisture Detector for Missile Systems.—A moisture detector has been developed by the U. S. Army Engineer Research and Development Laboratories, Fort Belvoir, Va., for use at missile installations and other areas where undetected moisture might cause malfunction of vital instruments and equipment.

Called an "electrolytic hydrometer," the instrument is capable of detecting moisture in high pressure air up to 6000 psi. and also of measuring moisture content down to a free air dew point of − 100° F. It weighs 35 lb., and is 11 in. high, 12 in. wide, and 15 in. long.

The instrument consists basically of an electrolytic cell, an ammeter and a direct current power source with associated pneumatic and electrical controls for efficient operation.

In use, the electrolytic type moisture analyzer absorbs and simultaneously electrolyzes all moisture from a sample gas stream. The absorption and electrolysis are accomplished in a small diameter tube containing a pair of closely spaced platinum wires wound as a double helix on its inner surface and coated with a thin film of phosphorous pentoxide, a material with a strong affinity for moisture.

The equipment was built by Consolidated Electrodynamics Corp. and Beckman Instruments Inc., under contracts with the Laboratories. Analyzers now in use are with the Redstone missile systems.

Electronic Tape Converters.—Minneapolis-Honeywell's Electronic Data Processing division has announced the development of electronic converters that automatically translate records on magnetic tape from IBM and Univac language into Honeywell language. The IBM converter has been completed and is in operation at the division's service bureau at Wellesley, Mass., where the Univac converter will soon be in operation.

For the present, Honeywell plans to use the new machines on a service basis to assist those users of Honeywell EDP systems who wish to convert non-Honeywell tapes into Honeywell tapes.

The translations proceed at whatever speed the non-Honeywell tapes read the recorded data. The translations often result in substantial tape savings, with the ratio running

as high as two and one-half non-Honeywell tapes to one Honeywell tape.

The equipment operates as follows: Univac and IBM tape units read the information recorded on their tapes into the Honeywell converters (the Model 831 for Univac and the Model 832 for IBM). This equipment converts the non-Honeywell signals into Honeywell words and passes them through a standard Honeywell tape control unit into the central processor of a Honeywell 800. The Honeywell 800, assisted by one of several specially-written programs, translates the IBM or Univac code into Honeywell code and records the output on Honeywell magnetic tapes.

Magnetic Mathboard Simplifies Instruction.—Mathematics can be more effectively presented to students with the impact gained from visualization with the new Magnetic Mathboard now available from Atomic Laboratories, Inc.

The Mathboard is a 3 ft. by 4 ft. magnetic blackboard that can be set on its own easel or on the chalk tray of a regular classroom blackboard. By means of white and black plastic geometric figures magnetized to the board, scores of mathematical functions can be visually illustrated by moving them around the board. Movable angle pieces, for example, show that the sum of a triangle's three angles equals 180 degrees.

In arithmetic, the Magnetic Mathboard illustrates slide rule operations of addition and subtraction, fractions and their manipulation; in algebra, it shows algebraic addition and subtraction, monomial multiplications, factoring the difference of two squares, graph work, arithmetic progression; in geometry, it illustrates the sum of angles in geometric shapes and other geometric functions; and in trigonometry, it shows multiplication and division using logarithms. A 26-page, fully illustrated manual provides a complete selection of demonstrations.

In addition to holding the various magnetized plastic parts, geometric shapes, numbers, letters, etc., the board functions as a regular blackboard and can be written on with chalk. The board comes with a regulation blackboard finish and is framed in aluminum trim.

Data-Processing System for Banking.—Burroughs Corporation has announced a new electronic system for speeding and controlling the interchange of more than 50 million checks every day among U. S. banks. Called the Burroughs B270 Electronic Data Processing System, it represents another major advance in bank automation equipment developed to curb rising operating costs and to alleviate a shortage of trained personnel.

Designed to provide automatic proof and transit operations, the B270 will also convert "on-us" items to magnetic tapes for subsequent computer processing and may be used for automatic deposit analysis as well as account reconciliation services.

The system's versatility will enable banks to expand their correspondent services and improve item control for either conventional bookkeeping or electronic data-processing systems, according to the company.

With ever growing volumes of checks flowing through their proof and transit departments, many banks are finding it increasingly difficult to maintain adequate staffs of well qualified proof clerks.

It takes several months to train a skilled operator of conventional proof machines, with much of this time

devoted to teaching complicated routing patterns. Automatic routing on the B270 eliminates this problem.

Furthermore, the system offers an answer to the profit squeeze caused by rising proof and transit costs through reduced item handling and greater overall efficiency and accuracy.

The B270 combines magnetic ink character recognition (MICR)—banking's "common machine language"—with high speed document sorting, multiple-list printing and solid-state electronics. This results in broad item distributions with a minimum of handling.

The system utilizes the world's fastest electronic sorter-reader, which reads checks and deposit tickets encoded with MICR at speeds up to 1560 items per minute, for input to the system. The sorter is also used to accomplish the physical sorting of all items.

Other units include an internally-programmed central processor with magnetic core memory, two ultra-high speed multiple-tape listers and a punch card reader for program loading and entry of miscellaneous information into the system.

Magnetic tape units are added to the system to permit conversion of "on-us" information for subsequent use in demand deposit data processing. These units also facilitate the use of the B270 as an off-line converter for various data processing systems.

Improved item control stems from the system's ability to produce master listings and detailed cash letters in which each item is identified with either the payor bank or the individual payor. This extended audit trail is obtained automatically, with no reduction in handling speed.

Application concepts of the B270 have been proven in Burroughs installations of magnetic document processing systems now in operaion at Michigan National Bank of Lansing, First Pennsylvania Banking & Trust Co. of Philadelphia and the Chicago Federal Reserve Bank.

Epoxy Silver Solder.—A silver conductive epoxy with a resistivity approaching the resistance of metals—between 0.01 ohms-centimeter and 0.0001 ohms-centimeter—has been developed by Epoxy Products, Inc., a division of Joseph Waldman & Sons, 137 Coit Street, Irvington, N. J. Epoxy solder has two principal advantages over conventional solder or brazing. It can be used at low temperature for components which are heat sensitive and it has no flux or residue to contaminate sensitive electronic components.

Epoxy solder is available in two paste forms—a one component heat curing (as low as 125° C.) paste and a two component room temperature cure paste. It is ideal for field repair work, particularly on crowded printed circuits where solder temperatures might damage heat-sensitive components. Its bond strength is superior to usual metal solders and it can be used to make a conductive bond with practically any material. Shear strength of a steel-to-steel bond is 3200 psi. Aluminum has approximately the same shear strength but is difficult to measure because the aluminum begins to tear before the bond fails.

When the bond hardens it can be planted with conventional plating techniques. Soldering can then be performed on the plated surface if desired.

Unlike conventional soldering and brazing, epoxy solder requires no flux and leaves no residue which might contaminate sensitive electronic components. Typical applications in-

clude bonding electronic tube caps to the glass tube, bonding the tantalum slug to the can of a tantalum capacitor, connections made to heat-sensitive diodes and transistors and grounding aluminum capacitors to steel chassis.

Conductive epoxies are intended primarily for replacing soldering and brazing operations which require high temperatures. When used with dissimilar metals, epoxy solder offers a bonding technique where no type of soldering or brazing would be possible.

Solid-State Crystal Oscillator.—A miniature, all solid-state voltage-controlled crystal oscillator (VCXO), the first standard model of a new line of similar products, was shown by Itek Electro-Products Company at the 1961 IRE Show, held in New York recently.

The Itek development gives radio engineers, for the first time, a crystal-controlled FM oscillator which can be directly frequency-modulated over a range as great as ± 0.2 per cent of center frequency, with excellent linearity and with virtually no degradation of crystal frequency stability over a wide range of temperatures. The units are expected to find wide application for FM generation, automatic frequency control and other operations in communications, telemetry, and guided missile systems.

VXCO Model 10M-WA has a center frequency of 10.7 mc and a deviation of ± 20 kc; linear to within ± 200 cps. Measured temperature drift is less than 1 kc. from − 40 to + 65° C. Output power is 5 mw.; size of the transistorized unit is only ¾ in. × 2 in. × 2½ in. over-all.

The standard Model 10M-WA is designed with a center frequency of 10.7 mc. to be compatible with Itek Electro-Products' 10.7 mc. Series of high-selectivity Hermes crystal filters and matching crystal discriminators. The 10M-WA is available in limited quantities for immediate delivery; other VCXO's, in the center frequency range 10 kc. to 30 mc. can be built to special order to meet both commercial and military system requirements. Units can also be supplied housed in temperature-controlled ovens which minimize temperature drift and increase frequency stability.

The VCXO circuit incorporates a unique embodiment of crystal resonator and voltage-variable reactance to permit direct frequency modulation of the oscillator. This technical achievement derives from a thorough mathematical analysis, by Itek Electro-Products' engineers, of the properties of piezoelectric crystals used with variable reactance networks of various types. The device is patented. It is claimed to be no more complex in size and number of components than conventional amplitude modulating systems, permitting the advantages of frequency modulation in many crystal-controlled applications where it had not been feasible before. This is especially true of portable equipment, where the conventional FM technique of multi-stage phase modulation synthesis requires considerably more complex circuitry and larger equipments.